DOCUMENT RESUME

ED 054 855 PS 004 911

AUTHOR	Rohwer, William D., Jr.; Ammon, Paul R.
TITLE	The Assessment and Improvement of Learning and Language Skills in Four and Five Year Old Culturally Disadvantaged Children. Final Report.
INSTITUTION	California Univ., Berkeley.
SPONS AGENCY	Office of Economic Opportunity, Washington, D.C.
PUB DATE	Jun 71
NOTE	140p.
EDRS PRICE	MF-$0.65 HC-$6.58
DESCRIPTORS	Caucasians; *Culturally Disadvantaged; Grade 1; Grade 3; Imitation; *Kindergarten Children; *Language Skills; *Learning; Lower Class; Middle Class; Negroes; *Preschool Children; Sentences; Testing; Vocabulary Development
IDENTIFIERS	Peabody Picture Vocabulary Test; PPVT

ABSTRACT

The objectives of this project were to assess and improve learning skills and language abilities in culturally disadvantaged children. The five studies included under this cover were prepared as self-contained reports yet they shared these common objectives. The studies are: (1) Learning Efficiency and Elaboration Training among Four- and Five-Year-Old Children; (2) Elaboration Instructions and Population Effects; (3) Individual Differences in the Learning of Verbally and Pictorially Presented Paired Associates; (4) Effects of Training in Vocabulary vs. Sentence Construction; and (5) Toward a Study of Sentence Imitation as a Technique for Assessing Language Ability. Appendixes give samples of language activities, lessons, and tests. (Author/AJ)

ED0 54855

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

THE ASSESSMENT AND IMPROVEMENT OF LEARNING AND LANGUAGE SKILLS IN FOUR AND FIVE YEAR OLD CULTURALLY DISADVANTAGED CHILDREN

William D. Rohwer, Jr.

Paul R. Ammon

University of California, Berkeley

FINAL REPORT

Office of Economic Opportunity

Contract Number B99-4776

June, 1971

PS 004911

Table of Contents

Page

List of Tables ii

List of Figures iv

List of Appendices v

Preface vi

Learning Efficiency and Elaboration Training among Four- and Five-Year-Old Children 1
William D. Rohwer, Jr., Mary Sue Ammon, and Joel R. Levin

Elaboration Instructions and Population Effects 20
William D. Rohwer, Jr., and Barbara J. Gerdes

Individual Differences in the Learning of Verbally and Pictorially Presented Paired Associates 29
Joel R. Levin, William D. Rohwer, Jr., and T. Anne Cleary

Effects of Training in Vocabulary vs. Sentence Construction 43
Paul R. Ammon and Mary Sue Ammon

Toward a Study of Sentence Imitation as a Technique for Assessing Language Ability 55

References 82

Appendices A-1

List of Tables

Table No. Page

1 Mean Chronological Age, Peabody Picture Vocabulary Test Mental Age and IQ, and Mean Number of Correct Responses on the Paired-Associate Pretest for the Initial Samples as a Function of Grade and Population (n=48) 8

2 Estimates of the Reliability of the Paired-Associate Pretest and of its Correlation with the Peabody Picture Vocabulary Test as a Function of Grades and Population in the Initial Samples 12

3 Correlations of PPVT and Paired-Associate Test with the Metropolitan Readiness Test as a function of Population in the initial kindergarten samples 14

4 Mean Number of Correct Responses (max.=25) on the Paired-Associate Posttest as a function of Grade, Population and Treatment 15

5 Mean Number of Correct Responses per Trial (max.=25) as a function of List, Population and Treatment 24

6 Mean Clustering Score (Z) per Trial as a function of List, Population and Treatment 26

7 Classifying List Performance of Hi and Lo P-V Types on V and P Items 33

8 Criterion List Performance of Hi and Lo P-V Types on V and P Items 35

9 Means and Standard Deviations of Form A Verbal (V) and Pictorial (P) Items for the Six Samples (48 Ss per sample) 37

10 Verbal and Pictorial Item-Type Correlations for the Six Samples 39

11 The Reliability of Raw and Standardized Differences between P and V Items 40

12 Assignment of Subjects in each Treatment Group to Experimenters A, B, and C 46

13 Preschool Results: Group Means on Five Dependent Measures from Posttest (23 Ss per group) 49

14 Kindergarten Results: Group Means on Three Dependent Measures from Posttest (22 Ss per group) 51

List of Tables (Cont'd)

Table No.		Page
15	List of Items from the Sentence Imitation Test	58
16	Item Numbers of 32 Sentences Representing Each Combination of Basic Sentence-Type X Kind of Elaboration	60
17	Item Numbers of 12 Sentences Derived from Two SVO Clauses	61
18	Means and Standard Deviations in Six Response Categories for Pretest Administration of the Sentence Imitation Test	65
19	Means and Standard Deviations for Three Variables from the Sentence Imitation Test (Pre- and Posttest)	66
20	Intercorrelations of SIT Variables for Preschool (Below Diagonal) and Kindergarten (Above Diagonal)	68
21	Means and Standard Deviations for Five Variables from Picture Interview for Preschool Subjects (N=12)	77
22	Intercorrelations of Five Variables from the Picture Interview, Pretest (Above Diagonal) and Posttest (Below Diagonal)	79
23	Intercorrelations of Sentence Imitation and Peabody Picture Vocabulary Scores with Five Variables from Picture Interview Pretest	81

List of Figures

Figure No. Page

1 Mean number of correct responses on the paired-associate pretest and reported by Rohwer, Ammon, Suzuki and Levin (1971) as a function of Population, Grade, and Item Type 10

2 Mean number of correct responses on the paired-associate posttest as a function of Population, Grade and Item Type 17

List of Appendices

A. Objectives and Procedures of Daily Treatment Sessions

B. List of Target Words for Vocabulary Training

C. List of Elaborative Constructions used for Sentence Training

D. Sample Lesson Plans for Language Training Sessions

E. Instructions for Transcribing from Tapes

F. Scoring System for the Sentence Imitation Test

G. Scoring System for the Picture Interview (With IBM Card Format)

Preface

This report is the product of research completed in contract between the University of California and the United States Office of Economic Opportunity. The project consisted of two major parts, one, under the direction of William D. Rohwer, Jr., concerned with assessing and improving learning skills in culturally disadvantaged children, and the second, under the direction of Paul R. Ammon, with the objective of assessing and improving language abilities in children drawn from the same population. The first part pertaining to learning skills is a direct extension of research completed by Rohwer in a previous contract between the University and the Office of Economic Opportunity (Jensen & Rohwer, 1970). Given the aim of improving the chances that the fate of culturally disadvantaged children will be a productive one, an exclusive emphasis on learning skills seemed patently inadequate; thus, Ammon's emphasis on the development of language abilities has invested the present project with substantial import.

Although it is already apparent that learning abilities and language abilities are closely related in the development of the child, an interweaving of the two parts of the project did not seem promising in view of the present status of theory and methodology in these problem areas. The distinctive character of the guiding theoretical notions and of the methods used for verifying hypotheses about learning skills and language abilities is reflected in the structure of the present report. Each of the five papers included has been prepared as a self-contained report and those responsible for each have been explicitly indicated by authorship. In our view, this collection represents progress both toward our separate and toward our common objectives--we regard the task of integration as one for the future.

We wish to recognize our dependence on those who have collaborated with us in completing the project by declaring our gratitude for their extraordinary competence and diligence. A number of persons filled the several roles necessary for conducting the studies under Rohwer's direction. Mrs. Mary Sue Ammon created the instructional materials and procedures for the study of elaboration training; in addition, she supervised the testing and treatment sessions, managing the severely complicated logistics required to bring this study to fruition. Dr. Joel Levin designed and supervised the methods and procedures of data analysis and tabulation for evaluating the training study; Dr. Nancy Suzuki performed the analyses of pretest data. All of the pretests, posttests, treatment sessions, data collection and tabulation were done by Miss Barbara Gerdes, Miss Carole King, Miss Caryl Sutton, and Miss Carolyn Saarni. Miss Gerdes also collected and analyzed the data for the study of free recall learning and collaborated in preparing the report; Miss Kathleen Woode provided consultation on the problems and methods used for estimating clustering in the free recall task.

Some of the same people worked with Ammon in the research on language. Mrs. Mary Sue Ammon's contribution to virtually every phase of the project was substantial and indispensable. Drs. Joel Levin and Nancy Suzuki—who know the items in the Sentence Imitation Test better than anyone else on earth—were most helpful in analyzing the results of the language training study. All of the raw data were collected by Mrs. Sharon Alexander, Mrs. Carol Meredith, Miss Sharon Rose, and Miss Caryl Sutton. In addition, Mrs. Ammon, Mrs. Meredith, and Miss Rose planned and conducted the training sessions. The task of transcribing literally miles of tape recordings was managed by Mrs. Phyllis Koppelman and Miss Patricia Stohr. Miss Sylvia Zuck skillfully double-checked the interview transcripts and conducted the word counts which were analyzed in the training study. Last but not least, Mrs. Judy Harker assisted Mrs. Ammon in getting the bugs out of the interview coding system and was responsible for the first analyses in which the system was applied.

Mrs. Carol Rohwer managed all of the administrative functions entailed by the entire project, as well as typing and compiling all of the manuscripts; her skill and persistence were responsible for the smooth maintenance of the enterprise and for the excellence of production in the final report. Miss Katherine Eardley prepared the figures and Mrs. Enid Goetzl proofread the copy.

We have no adequate means for paying our debt to the administrators, teachers and children in the schools where the research was conducted. They tolerated our intrusions and worked cooperatively with us to produce the information reported here. Thus, we are enduringly grateful to hundreds of persons in the Bay Area cities of Berkeley, Lafayette, Moraga, Oakland, Orinda, and Richmond.

Learning Efficiency and Elaboration Training among Four- and Five-Year-Old Children[1]

William D. Rohwer, Jr., Mary Sue Ammon and Joel R. Levin

The efficiency of paired-associate learning in young children varies markedly with presentation conditions. For cases where the task is that of learning lists of noun pairs, recent research has identified particular ways of presenting the nouns that control the rate of acquisition. Considering the method of aural presentation as a baseline, each of the following variants is associated with greater and greater increments in learning efficiency (Rohwer, Ammon, Suzuki & Levin, 1971): pictures (i.e., photographs of the two objects denoted by each pair of nouns); pictures plus nouns; pictures plus an aurally presented sentence containing the two nouns; action pictures (i.e., motion picture sequences displaying a brief episode involving the two objects by the nouns) plus nouns. The power of two of the four methods, sentence context and action pictures, to increase learning efficiency has been demonstrated in numerous experiments with children (e.g., Rohwer, 1967, Experiments I-XIII; Rohwer, Ammon, Suzuki & Levin, 1971).

Rohwer (1970) has proposed an hypothesis to account for the effects of these presentation methods on paired-associate learning. It is assumed that learning any noun pair consists of modifying the initial meaning of the two nouns such that they share at least one, unspecified, semantic aspect. This process of semantic modification is referred to by the term 'elaboration.' The methods of presentation mentioned, from aural nouns to action pictures, are thought to constitute increasingly potent prompts for elaborative activity. Thus, the probability that elaborative activity will be evoked is greater when nouns are presented in the form of their object referents than when presented as spoken words.

Although the success of these presentation methods suggests possible educational applications in order to improve learning efficiency, it is of interest to explore another means of achieving this same goal. In particular, the present study was undertaken to determine the feasibility of inducing children themselves to supply prompts to elaborative activity when faced with the task of noun-pair learning. With respect to educational utility, this approach has a distinct advantage over that of incorporating elaborative prompts in instructional materials, namely, that the child himself can generate the necessary conditions for efficient learning rather than relying on the foresight and skill of others.

---

[1] The conduct of this study literally would have been impossible without the assistance and cooperation of many persons. Barbara Gerdes, Carol King, Carolyn Saarni, and Caryl Sutton with skill and impressive endurance conducted all of the pretest, treatment, and posttest sessions in addition to tabulating and keypunching all of the data. Dr. Nancy Suzuki performed many of the analyses reported here. The principals, directors, teachers and children in all of the schools not only tolerated but helped us to complete the study. We gratefully acknowledge our debt to all.

The promise of the approach envisioned here is suggested by the success of some prior studies. Jensen and Rohwer (1965) observed marked facilitation of noun-pair learning as a result of directing second-, fourth-, and sixth-grade children to construct and utter sentences containing each pair of nouns immediately before performing the paired-associate task. Rohwer (1967) has reported similar results for samples of kindergarten, first- and third-grade children. In a study that more closely approximates an instructional methodology, Milgram (1967a)also found substantial positive effects of inducing subject-generated prompts to elaborative activity. A multiple-list training method was used in which each child was asked to learn four successive lists of noun pairs. In the treatment condition, the pairs were presented in sentence contexts on the first list, subjects themselves were instructed to generate the sentence contexts on the second list, and standard paired-associate learning instructions were given on lists three and four. The purpose of the third list was to assess transfer relative to a control group that had also learned two previous lists but under standard instructions. The fourth list was administered a week after the third to permit a determination of the degree of retention of the elaboration set presumably induced by the conditions of learning the first two lists. On both lists three and four, performance among treatment subjects was markedly more efficient than among control subjects, indicating that this method was successful in inducing the behavior of generating elaborative prompts in the absence of external stimulation to do so.

The immediate precursor of the present study was an experiment reported by Rohwer and Ammon (in press). Samples of second-grade children were provided with five half-hour sessions on successive school days in which they either practiced learning lists of noun pairs (Practice) or received instruction in generating elaborative prompts (Training). An additional factor in the design was Population; half the children were drawn from a low-SES black residential area and the other half from a high-SES white residential area. Treatment effects, as measured on a posttest that consisted of a 25-item list of noun pairs, revealed greater learning efficiency in the Training condition than in the Practice and Control conditions among the low-SES black children; for the high-SES white children, performance in the Practice condition was superior to that in the Control and as good as that in the Training condition. Accordingly, it was concluded that the training method used was of sufficient promise to warrant further experimentation.

The present study was an attempt to extend the work reported by Rohwer and Ammon (in press) in two directions: first, to augment the elaboration training and to compare two methods of offering it to children; and, second, to sample from younger age levels. This second modification deserves additional comment since the decision relates to sampling from two populations--high-SES white and low-SES black. Previous comparisons of performance on paired-associate tasks in these two populations have yielded significant differences favoring high-SES white children only for children of pre-school and kindergarten ages (Semler & Iscoe, 1963; Rohwer, 1967; Rohwer, Ammon, Suzuki & Levin, 1971) whereas equivalent levels of performance are usually observed for children in the first through sixth grades of elementary school. These results suggested that elaboration training might have more pronounced effects among four- and five-year olds than among older children. If so, it was reasoned that instruction in the skills of producing elaborative prompts might be recommended for inclusion in curricula for preschool programs. Finally, the rationale for including children from both populations in the present study was to permit an assessment of the generality of those training effects that might emerge.

Method

Subjects. Four samples were drawn. Two of these consisted of children enrolled in kindergarten classes in elementary schools and the other two of children enrolled in pre-schools. In the case of the pre-school children, sampling was limited to those parts of the populations where the children were eligible, in terms of chronological age, for entrance into public school kindergarten at the beginning of the Fall term following the academic year in which the study was conducted. At each grade level, half the children were drawn from schools serving low-SES black residential areas and the other half from schools serving high-SES white residential areas as indexed by average census tract data collected in the 1960 survey. In the case of the low-SES black samples, the preschool children were drawn from the same schools (i.e., public schools) as the kindergarten children. In contrast, the high-SES white preschool children were drawn from private nursery schools. In order to insure sampling from approximately the same population as for the high-SES white kindergarten children, these nursery schools were selected in residential areas which were either the same as or adjacent to those served by the public school kindergartens.

With the intent of retaining a sample size of 48 children in each of the four goups at the end of the study, 72 children were initially selected at random from the four populations. Of this total, 48 were designated for initial inclusion and an additional 24 as alternates. The chronological ages of those included initially are presented in Table 1 as a function of grade and population. (Table 1 may be found on page 8.)

Design. In a 2x2x6x4 design the factors were: Grades, Population, Levels, and Treatment. The first two factors provided comparisons of pre-school and kindergarten children, and of high-SES white and low-SES black children, respectively. Within each of the four groups defined by the first two factors, subjects were assigned in equal numbers to one or another of six levels defined by performance on a pretest comprised of two paired-associate lists of 25 noun pairs. Thus, within a group, a total of 12 subjects was assigned to each level, although only 8 were initially designated for inclusion in the study; the remaining four were alternates. Of the 12 at each level, three were randomly assigned to each of the treatment conditions.

The four treatment conditions, tutorial, didactic, practice and control, were distinguished by the character of events that intervened between the paired associate pretests and posttests. In all but the control condition, these events occurred in the context of six daily sessions for every child of approximately 15 to 20 minutes each. In the tutorial and didactic conditions, the objective of the sessions was to provide training in skills for producing two kinds of prompts to elaborative activity in connection with the task of learning noun pairs: (a) the construction of sentences containing the two nouns in each pair; and (b) envisioning the two objects denoted by each noun pair in a scene where they are either interrelated or interacting. The sequence of activities in the training sessions was designed to include the following steps: an illustration of the difficulty of learning paired associates without elaborative prompts; a description of the two prompting techniques with examples; instructions to the child to produce the two kinds of prompts; practice at producing the two kinds of prompts with a variety of materials (words, pictures, objects, cut-outs, etc.); practice at producing the two kinds of prompts when pairs are presented at fixed pacing intervals; and gradual withdrawal of the direction to produce the prompts. Examples of the activities used to accomplish these steps include: presenting eight pairs of familiar objects to the child, asking him to construct a story about the two objects in each pair, then to act out the story and tell it verbally as well; presenting pairs of objects by videotape, asking the child to envision an interaction for each pair, showing such a filmed interaction, and asking the child to describe the scene verbally.

There were two major differences between the tutorial and didactic conditions. In the tutorial treatment, the child was required to demonstrate (act out, describe verbally) the elaborative prompts he was asked to construct and the experimenter provided feedback about the adequacy of the prompt. In the didactic treatment, examples of elaborative prompts were demonstrated by the experimenter and the child was instructed to generate such prompts but he was not asked to communicate them to the experimenter so that no feedback was given about their adequacy. The purpose of this manipulation was to determine the effectiveness of providing training without directly monitoring the child's acquisition of the skills taught. If the didactic treatment proved to be as effective as the tutorial, it was reasoned that the feasibility of implementing elaborative-prompt training would be enhanced since it could be provided inexpensively, either by machine or in small groups by a teacher.

The objective of the daily sessions in the practice condition was to control for the generalized positive transfer presumed to accrue in the tutorial and didactic conditions from learning the several lists used to illustrate elaborative prompts and for possible effects of contact with an individual experimenter during the daily sessions. Accordingly, the materials used in the practice condition were identical with those used in the tutorial and didactic conditions; all children were provided with the same amount of practice in learning lists of paired associates. In addition a number of other procedures were constant across the three conditions: the use of puppets, performance charts, encouraging comments, etc. as incentives for learning. In summary, the structure and content of sessions in the practice condition were designed to differ from the tutorial and didactic conditions only with respect to direct instruction in techniques of elaborative prompting. A detailed account of the various sessions is presented in Appendix A.

It was also planned to compare the practice and control conditions to permit an assessment of the effects of all of the activities in the training sessions except instruction in elaborative prompting. Thus, all children assigned to the control condition received the pretests and posttests but no daily sessions with an experimenter.

Materials. Prior to the paired-associate pretest, all subjects were given the Peabody Picture Vocabulary Test (PPVT). The PPVT is described in detail elsewhere (Dunn, 1965). It consists of a series of items in each of which a page bearing pictures of four familiar objects or events is exposed to the subject while the experimenter utters a word that denotes the content of one of the pictures. The subject's task is to designate the picture denoted. The test was administered and scored in accord with instructions given in the manual, yielding a raw score, MA and IQ for each child.

The paired-associate pretest consisted of two 25-item mixed lists administered for a total of two pairing-test cycles each. Five different types of items were represented equally often in each list. Although the task for every item was the same, i.e., to learn a noun pair, the pairs were presented in five different ways: as orally uttered words (Nouns); as pictures of the objects denoted by the nouns (Pictures); as words and as pictures (Pictures + Nouns); as orally uttered nouns in the context of a sentence presented along with the pictures of denoted objects (Pictures + Sentences); and as nouns uttered while action (motion) pictures were presented depicting the two objects in an interaction (Action + Nouns). All of these materials were recorded on videotape and presented by means of playback through a television monitor. During the pairing trials, the items were presented at a 4-sec. rate and were ordered so that every item type was represented once in each successive set of five items; the order of item types within a set was random.

During the test trials, the stimulus members of the pairs were presented at a 4-sec. rate and the subject's task was to utter the name of the response member during the interval. These materials were also recorded on videotape and played back through the monitor. As each stimulus item was presented, the audio portion of the tape played a question, "What went with the (stimulus noun)," whereupon the child was to utter the appropriate response noun. This question was inserted to encourage the children to respond to every item; in this respect the paired-associate test differed from the version used previously (Rohwer, Ammon, Suzuki & Levin, 1971; Rohwer & Ammon, in press). A different random order of pairs was used on each pairing trial and other random orders of stimulus terms were used on the test trials.

The instructions contained a description of the task and asked the children to learn the pairs so as to be able to produce the response term for each stimulus term presented on the test trials. The instructions were clarified by the presentation of five sample items, one of each

PS 004911

type, repeatedly, until the subject attained a criterion of three correct responses on test trials.

The paired-associate posttest also consisted of two 25-item mixed lists of noun pairs in which the format and characteristics of the lists were identical to those described for the pretest. The difference was that the lists were composed entirely of nouns other than those appearing on the pretest lists. All nouns were of high familiarity, as were their object referents, and they were paired so as to avoid obvious associations.

The materials for the training sessions have already been alluded to; they included puppets, objects, pictures, videotapes, audiotapes, stickers, etc. The various items are enumerated in Appendix A.

Procedure. Every child was scheduled for two initial sessions and a final session. During the first of the initial sessions, the PPVT and the first list of the paired-associate pretest were administered; during the second, typically one or two days later, the second list of the paired-associate pretest was administered. The final session consisted of the administration of the two paired-associate posttest lists. The interval between the pretest and the posttest varied from 15 to 30 days depending on scheduling constraints. In virtually all cases of children assigned to one or another of the contact treatment conditions, the interval between the last treatment session and the posttest session ranged from two to five days. The six treatment sessions were scheduled for each child over a two-week period to permit, whenever possible, the administration of the posttest at the end of the second week. In the cases of absence from school, make-up sessions were arranged except when the absence was extended beyond two or three days; in the latter instances, the absentee was dropped from the sample and replaced by an available alternate. All procedures, pretest, posttest, and treatment were conducted with the children individually.

Four white female experimenters conducted the study. Each was involved equally often in all phases--pretest, posttest and treatment sessions--and, in the case of treatment sessions, equally often in each of the three contact conditions. Once the assignment of a particular child to a particular experimenter was made for the treatment sessions, all of that child's sessions were conducted by that experimenter. However, no experimenter conducted more than one phase of the study with a given child. Thus, if an experimenter administered the pretest to a child, she did not administer the treatment sessions nor the posttest. The purpose of this complicated scheduling precaution was to assure, as much as possible, an evaluation of the treatment effects unbiased by other information about the child assessed.

Results

To be presented here are the results pertaining to four kinds of data yielded by the present study: PPVT scores for the initial sample; paired-associate pretest outcomes for the initial sample; intertask relationships; and, finally, the outcome of the training study itself as indexed by performance on the paired-associate posttest.

Peabody Picture Vocabulary Test. The results for the four initial samples ($n = 48$) on the PPVT with respect to MA and IQ are presented in Table 1. Scores on the MA variable were subjected to analysis of

- - - - - - - - - - - - - -

Insert Table 1 about here

- - - - - - - - - - - - - -

variance in which the design was a three-way factorial. The factors of Population and Sex were treated as nested in Grades. As an inspection of Table 1 suggests, the main effect of Grades was significant ($F = 87.91$, $df = 1/184$, $p < .05$), such that the average MA of kindergarten children was higher than that for preschool subjects. The Population effect was significant at both grade levels ($Fs = 38.37, 54.95$, $dfs = 1/184$, $ps < .05$, for preschool and kindergarten respectively); in each case the mean MA of the high-SES white children was above that of the low-SES black children. The effect associated with Sex was significant only in the kindergarten samples ($F = 14.00$, $df = 1/184$, $p < .05$) not among the preschool children ($F < 1$). The mean MA of the male kindergarten subjects was higher than that of the females (75.04 vs. 65.19 mos.). The interaction of Population and Sex was not significant for either preschool ($F = 1.92$, $df = 1/184$, $p > .05$) or for kindergarten samples ($F < 1$). In summary, the results on the PPVT showed a marked advantage favoring the high-SES white samples; indeed the proportion of the total sums of squares associated with the Population effect was .24.

Paired-Associate Pretest. Performance on the paired-associate pretest was indexed in terms of the mean number of correct responses given on the two test trials of each list. With respect to between-subjects effects, the analysis of variance design applied to these data was the same as that described for the PPVT. The relevant means are presented in Table 1. Performance among kindergarten children was significantly better than among preschool children ($F = 46.33$, $df = 1/184$, $p < .05$). The Population effect was also significant at both grade levels ($Fs = 17.07, 22.42$, $dfs = 1/184$, $ps < .05$, for preschool and kindergarten respectively) favoring the high-SES white children in each sample. Of the total between-subjects sums of squares, 14% was associated with the main effects of Population. This outcome is in accord with those of previous studies that have included samples of children from these two populations (high-SES white and low-SES black) drawn from preschool and kindergarten ages (Rohwer, 1967, Experiments XII and XIII; Rohwer, Ammon, Suzuki & Levin, 1971).

The effects associated with Sex were similar to those observed on the PPVT: boys produced more correct responses than girls at the kindergarten level ($F = 8.32$, $df = 1/184$, $p < .05$) but the difference was not significant among preschool children ($F < 1$). The Population x Sex interaction was significant at preschool ($F = 7.08$, $df = 1/184$, $p < .05$) but not at kindergarten ($F < 1$). Descriptively, the interaction in the preschool samples apparently emerges because the boys produced more correct responses than the girls among the high-SES white children whereas the girls performed better than the boys in the low-SES black sample.

Table 1

Mean Chronological Age, Peabody Picture Vocabulary Test Mental Age and IQ, and Mean Number of Correct Responses on the Paired-Associate Pretest for the Initial Samples as a Function of Grade and Population (n=48)

Grade	Population	CA (mos.)	MA (mos.)	IQ	Number Correct (max. = 25)
Preschool	High-SES white	52.98	60.81	106.79	6.97
	Low-SES black	54.98	44.50	82.42	4.37
Kindergarten	High-SES white	66.54	79.88	110.83	10.19
	Low-SES black	67.81	60.35	89.79	7.21

Both of the dependent variables examined thus far, PPVT MA and total score on the paired-associate pretest, have revealed a consistent result: low-SES black children do not perform at levels as high as those achieved by high-SES white children.

In the case of the paired-associate pretest, the Population difference can be analyzed further by examining the results as a function of Item Types. For this purpose, the factors of Lists, Trials and Item Types were included in a repeated measures analysis of variance; this analysis represents the within-subjects component of the design previously described. All subjects received two 25-item lists containing five types of items, so that the maximum score for each item type, averaged across lists and trials, was five. The results are presented in Figure 1 as a function of Grade and Population. The main effect of Item Types was significant at both grade levels (*Fs* = 140.12, 268.72, *dfs* = 4/736, *ps* < .05, for preschool and kindergarten respectively). Comparisons among types of items were made by means of the Scheffé method with the probability of a Type I error set at .05. These contrasts revealed the same pattern of effects at both grade levels: Action + Nouns items were superior to Pictures + Sentences and Pictures + Nouns which, in turn, were superior to Pictures and to Nouns items. None of the other adjacent pair-wise differences was significant. This outcome contrasts sharply with previous results obtained in two specific respects; Rohwer, Ammon, Suzuki & Levin (1971) reported significant differences between Pictures and Nouns as well as between Pictures + Sentences and Pictures + Nouns. Since the only methodological difference between their study and the present one was the use of the test trial prompt, "What went with the (stimulus noun)", the discrepancy may be due to that change.

- - - - - - - - - - - - - - -

Insert Figure 1 about here

- - - - - - - - - - - - - - -

Since the Population x Item Types interaction was significant at both grade levels (*Fs* = 8.24, 6.30, *dfs* = 4/736, *ps* < .05, for preschool and kindergarten respectively) tests were made of the simple effects of Population within each Item Type to locate more precisely the difference in performance between high-SES white and low-SES black subjects. In the preschool samples, this analysis revealed significant Population effects for the three types of items: Pictures + Nouns, Pictures + Sentences, and Action + Nouns (*Fs* = 45.13, 42.11, 45.75, *dfs* = 1/736, *ps* < .05). Among the kindergarten children, these same three kinds of items yielded significant Population effects (*Fs* = 44.52, 24.36, 69.18, *dfs* = 1/736, *ps* < .05) as did Nouns items (*F* = 21.73, *df* = 1/736, *p* < .05). The locus of the Population effects at the kindergarten level are at variance with results previously reported by Rohwer, Ammon, Suzuki & Levin (1971) in that Population differences were not as pronounced for Pictures + Nouns items as they are in the present study. The rightmost panel of Figure 1 permits a comparison of the two sets of results.

In summary, this detailed analysis of performance on the paired-associate pretest suggests: (a) that children from both populations do not make effective use of one kind of elaborative prompt, sentence context,

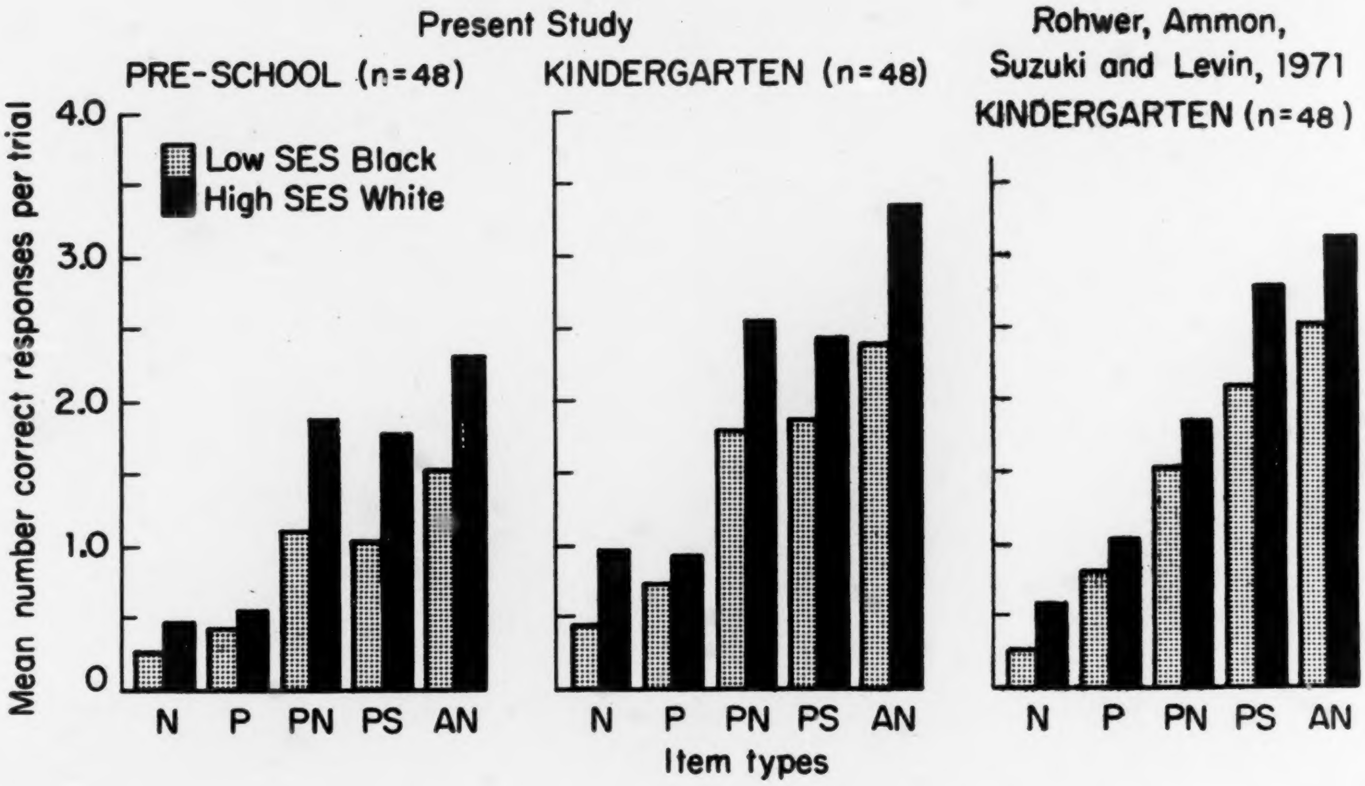

Figure 1. Mean number of correct responses on the paired-associate pretest and reported by Rohwer, Ammon, Suzuki and Levin (1971) as a function of Population, Grade, and Item Type.

known to increase learning efficiency; and (b) that relative to the high-SES white children, low-SES black children do not derive optimal benefit from presentation conditions designed to improve performance in noun-pair learning, namely, Pictures + Nouns, Pictures + Sentences, and Action + Nouns. Accordingly, training in the production and use of elaborative prompts appears quite promising for enhancing the learning proficiency of children from all of the samples selected.

Although a number of other within-subjects effects were significant (e.g., Trials, Lists, Item Types x Lists) none altered the conclusions drawn from the effects already described and none are germane to the purpose of the study. Thus, they will not be reported in detail.

Intertask Relationships. Reliability estimates of the paired-associate pretest can be made for the initial samples of the present study in terms of the correlations between performance on the two lists that comprised the test. These estimates are presented in Table 2 for scores on each Item Type separately and for total scores as a function of Population and Grade. The reliability estimates of the Item Types considered individually are very modest whereas those for the total score more nearly approximate an acceptable level, especially in view of the age of the children sampled and the relative novelty of the first list learned as compared with the second.

- - - - - - - - - - - - - -

Insert Table 2 about here

- - - - - - - - - - - - - -

Table 2 also displays the correlation coefficients between performance on the paired-associate pretest and PPVT IQ. The variable of IQ, rather than that of MA, was selected for examining these relationships since, in effect, it partials out the mutual correlation of performance on the two tasks with CA. An inspection of Table 2 reveals correlations between total score on the paired-associate test and IQ that are moderately high considering that the reliability of each measure appears to be approximately .75. Contrary to the hypothesis advanced by Jensen (1969) it is interesting to note that the correlation between IQ and paired-associate performance is of about the same magnitude for both high-SES white and low-SES black children. As for the correlations between separate Item Type scores and IQ, the general pattern appears to be that the relationship is more consistent for the three easier types than for the more difficult ones. This pattern may be attributable to a restriction on the range of scores yielded by the Nouns and by the Pictures types.

After the present study was completed, the Metropolitan Readiness Test was administered in the school districts from which the two kindergarten samples were drawn. It was possible to obtain scores from this test for 79 of the 96 children in the initial samples. To examine the relationships among the skills measured by the paired-associate test, the PPVT and the readiness test, the appropriate correlations were computed. These are displayed in Table 3 for PPVT, paired-associate total score and, in terms of multiple correlation coefficients, for the

Table 2

Estimates of the Reliability of the Paired-Associate Pretest
and of its Correlation with the Peabody Picture Vocabulary Test
as a Function of Grades and Population in the Initial Samples

Grade	Population	n	Product-Moment Correlation Coefficients											
			List 1 vs. List 2						PA vs. IQ					
			N	P	PN	PS	AN	Tot.	N	P	PN	PS	AN	Tot.
Preschool	High-SES white	48	$\underline{.43}$	.27	$\underline{.66}$	$\underline{.47}$	$\underline{.50}$	$\underline{.71}$	-.01	.27	$\underline{.41}$	$\underline{.45}$	$\underline{.39}$	$\underline{.41}$
	Low-SES black	48	$\underline{.41}$	.31	$\underline{.58}$	$\underline{.55}$	$\underline{.73}$	$\underline{.84}$	$\underline{.48}$	.11	$\underline{.52}$	$\underline{.44}$	$\underline{.45}$	$\underline{.50}$
Kindergarten	High-SES white	48	$\underline{.56}$	$\underline{.45}$	$\underline{.45}$	.23	$\underline{.40}$	$\underline{.78}$	.33	$\underline{.49}$	$\underline{.49}$	$\underline{.49}$	$\underline{.43}$	$\underline{.60}$
	Low-SES black	48	.31	.20	.33	$\underline{.49}$	$\underline{.62}$	$\underline{.68}$	.31	.28	$\underline{.50}$	$\underline{.52}$	$\underline{.44}$	$\underline{.52}$

Note— Underscored Coefficients: $p \leq .01$

- - - - - - - - - - - - - -
Insert Table 3 about here
- - - - - - - - - - - - - -

paired-associate Item Types yielded by entering subjects' scores separately for each type. Two features of these coefficients are noteworthy: first, the paired-associate test, especially when component scores are used, appears to measure some of the same skills as indexed by the readiness test; and second, there appears to be more overlap among these skills for the low-SES black children than for the high-SES white for whom the relationship with the readiness test seems stronger for the PPVT than for the paired-associate test. These results, coupled with those reported by Rohwer and Levin (1971) suggest that the paired-associate test taps school-relevant skills to a substantial degree among young children.

Elaboration Training. The results relevant to an assessment of the efficacy of training in elaborative prompts will be considered in two parts, one concerning the outcome as reflected in between-subjects effects and the other pertaining to within-subjects effects.

With regard to between-subjects effects, a four-way analysis of variance design was applied to the variable of total score on the paired-associate posttest. The factors were Population, Level, and Treatment, all nested in Grade. The results, averaged across lists and trials, are presented in Table 4. Also presented there are the final sample sizes in each cell. As these numbers indicate, attrition was not equivalent across conditions and groups; in some cases, the rate was disturbingly high. Accordingly, the data for all subjects tested and trained, including alternates was included for the analysis of training effects.

- - - - - - - - - - - - - -
Insert Table 4 about here
- - - - - - - - - - - - - -

The main effect of Grade was significant ($F = 29.20$, $df = 1/89$, $p < .05$) such that kindergarten children made more correct responses than preschool children. The main effect of Population was also significant, but only among the preschool children ($F = 29.91$, $df = 1/89$, $p < .05$). Although the main effect associated with the Population factor was not significant for the kindergarten samples ($F = 3.65$, $df = 1/89$, $p > .05$), the interaction, Population x Level, was ($F = 2.68$, $df = 5/89$, $p < .05$). Descriptively, the form of the interaction was such as to imply that a population difference in posttest performance was confined entirely to children who produced low scores on the pretest whereas no such difference was apparent for children who had produced high pretest scores. This interaction, however, was not significant at the preschool level ($F < 1$).

The main effects associated with Levels were significant for both preschool and kindergarten samples ($Fs = 10.20, 9.66$, $dfs = 5/89$, $ps < .05$). This outcome implies that the use of a levels design was successful in increasing the power of the tests for treatment effects. It also implies

Table 3

Correlations of PPVT and Paired-Associate test with the Metropolitan Readiness Test as a function of Population in the initial kindergarten samples

Population		PPVT	Paired-Associate Total	Item Types
	n	r	r	R
High-SES white	39	.56	.38	.45
Low-SES black	36	.47	.62	.68

Table 4

Mean Number of Correct Responses (max. = 25) on the Paired-Associate Posttest as a function of Grade, Population and Treatment

Grade	Population	Treatment									
		Tutorial		Didactic		Practice		Control		All	
		n	mean	n	mean	n	mean	n	mean	n	mean
Preschool	High-SES white	8	8.95	11	12.55	11	9.70	12	9.30	42	10.20
	Low-SES black	12	6.45	13	7.50	10	7.15	13	5.95	48	6.75
Kindergarten	High-SES white	12	11.05	11	12.20	10	12.05	12	11.55	45	11.70
	Low-SES black	12	11.10	12	10.45	13	10.00	11	10.10	48	10.40

substantial stability of individual differences in performance on the paired-associate task from pretest to posttest.

Tests for the effects of Treatment were of principal interest in view of the fact that the study was conducted to evaluate the efficacy of training in the production of elaborative prompts. The results were appraised by forming, and testing three orthogonal contrasts at each grade: tutorial vs. didactic, practice vs. control, and the average of tutorial and didactic vs. the average of practice and control. In the preschool samples, none of the three contrasts was significant (*F*s = 1.37, 1.37, 2.06, *dfs* = 1/89, *p*s > .05, respectively). Similarly, no significant treatment effects were detected for the kindergarten samples (all *F*s < 1). Furthermore, none of the interactions of Treatments with other factors were significant; in fact, the largest of the relevant *F* ratios was only 1.03. Thus, the results were unequivocal: as measured by total score on the paired-associate posttest, the study provides no evidence whatever that training in producing elaborative prompts improves paired-associate learning efficiency for any of the samples included in the study.

The within-subjects portion of the analysis treated the factors of Lists, Trials and Item Types as repeated measures. Because these factors, especially when combined with the between-subjects factors, provide so large a number of statistical tests, the probability of a Type I error for within-subjects effects was set at .01.

The main effect of Lists was significant (*F* = 14.47, *df* = 1/89, *p* < .01) indicating a difference in the difficulty of the two lists that formed the posttest. Trials was also significant (*F* = 685.96, *df* = 1/89, *p* < .01) such that more correct responses were made on the second than on the first trials of each list. An inspection of the interactions involving the factors of List and Trial with other factors in the design revealed no other significant effects.

The results concerning differences in performance associated with Item Types are displayed in Figure 2 averaged across Lists, Trials,

- - - - - - - - - - - - - - -

Insert Figure 2 about here

- - - - - - - - - - - - - - -

Treatments and Levels. An examination of these results indicates that the pattern of performance across Item Types is very similar to that observed by Rohwer, Ammon, Suzuki and Levin (1971). Thus, unlike the outcome on the pretest, the posttest data yield evidence of increasing learning efficiency from Nouns items to Action + Nouns items.

The main effect of Item Types was significant (*F* = 426.13, *df* = 4/356, *p* < .01) as were three of the numerous interactions involving Item Types. One of these, Item Types x Population for the kindergarten samples (*F* = 5.25, *df* = 4/356, *p* < .01), may be described as showing differences favoring the high-SES white children only on the Nouns, Pictures + Sentences, and Action + Nouns items. The interaction was not significant for the preschool samples (*F* = 2.20, *df* = 4/356, *p* > .01). The remaining

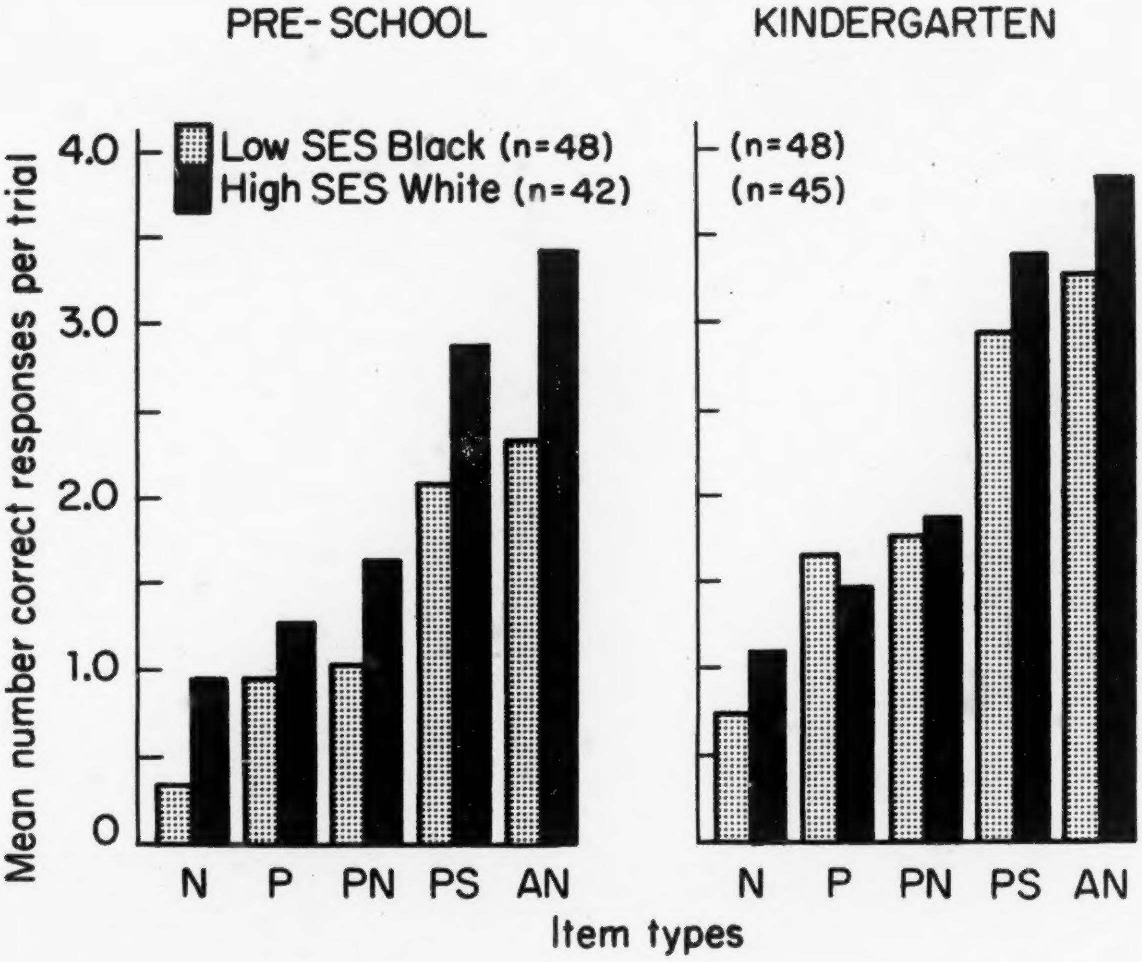

Figure 2. Mean number of correct responses on the paired-associate posttest as a function of Population, Grade and Item Type.

two interactions were higher-order ones, the complexity of which defy clear description. Since neither alters the conclusions implied by the other tests reported, they will not be presented here.

Discussion

With regard to the central purpose of the present study, the results were quite clear: training in elaborative prompts did not substantially improve the efficiency of paired-associate performance in the preschool and kindergarten children sampled. Although a null result of this kind is often uninformative about the validity of the hypothesis under test, in the present case it offers guidance. A treatment previously shown effective with samples of second-grade children has failed to produce performance differences in preschool and kindergarten children drawn from the same populations. In addition, even though the attempt to teach prompting techniques to children in the tutorial and didactic conditions was not successful, these same prompts were effective in promoting efficient learning when they were incorporated directly in the learning materials. Thus, one implication of the study is that children of the ages sampled derive more benefit from elaborative prompts presented with the items to be learned than from extensive instructional sessions designed to encourage the children to produce such prompts themselves.

If the discrepancy between the results of the present study and those reported previously by Rohwer and Ammon (in press) for second-grade children is attributable to differences in the ages of the subjects sampled, recent formulations drawing on a Piagetian framework provide a means of comprehending the phenomenon. Elkind (1969) and Kohlberg (1968) for example, have suggested that certain kinds of autonomous cognitive operations can be more effectively fostered at older than at the younger age levels treated in the present study. Similarly, Rohwer (in press) has argued that the capacity for engaging autonomously in the activities of elaborative prompting only begins to emerge in the preadolescent period, not during early childhood. Accordingly, it might be worthwhile to conduct an evaluation of the kinds of tutorial and didactic treatment used here among children aged ten to twelve rather than four and five.

Two other matters concerning the outcome of the present study deserve brief comment; both pertain to the paired-associate test used to estimate initial and final levels of performance. Both in terms of reliability estimates and in terms of relationships with other measures, specifically the Metropolitan Readiness Test, the method of mixed-list paired-associates appears very promising as a tool for analyzing individual differences in learning proficiency. The present results add further encouragement to the efforts reported here and elsewhere (Rohwer, Ammon, Suzuki & Levin, 1971; Rohwer & Levin, 1971; Levin, Rohwer & Cleary, 1971) to develop an instrument capable of yielding valid and reliable information about differences among children in the characteristics of their learning processes.

Finally, it should be reiterated that the Item Types effects detected on the pretest version of the paired-associate mixed list failed to replicate previous results (Rohwer, Ammon, Suzuki & Levin, 1971). The most obvious discrepancy was the fact that sentence contexts, (Pictures + Sentences) produced no more correct responses than simply supplying the names of the pictures (Pictures + Nouns). Since the sentence context effect has been shown repeatedly to be robust, it is of interest to determine why it did not emerge in the present study. The most likely explanation refers to the fact that in the present study, the time allotted for responding to each item on the test trial was severely reduced by adding the question, "What went with the (stimulus noun)?", to the presentation of every stimulus item. The practical effect of this change in procedure was to provide subjects only about 1 sec. for uttering the correct response noun in each pair. Support for the supposition that this modification was responsible for the observed discrepancy is offered by the results of another study (Rohwer & Ammon, 1968) in which test-trial pacing rate was experimentally manipulated. Rohwer and Ammon (1968) reported that sentence contexts significantly facilitated noun-pair learning when the test-trial rate was as slow as two or four seconds but not when it was as fast as one second. Thus, the sentence context effect is robust only within specifiable boundaries that may have been exceeded in the procedure used here for administering the paired-associate pretest.

Elaboration Instructions and Population Effects in Free Recall Learning[1]

William D. Rohwer, Jr. and Barbara J. Gerdes

The boundaries on an instructional or experimental effect are of critical importance, both for the development of psychological theory and for its application to education. The purpose of the present study was to assess the generality of one such effect with respect to a learning task other than the one with which the effect was initially discovered. The effect in question was first observed in a noun-pair learning task administered to children in accord with the method of paired-associates (Jensen & Rohwer, 1965). All children were asked to learn a list of noun pairs presented by means of pictures showing the objects denoted by the nouns. Those in the control condition were simply instructed to study the pairs whereas those in the treatment condition were instructed to generate a sentence or story containing the two nouns in each pair (e.g., the FOOT kicked the HOUSE). The sentence instructions resulted in performance that was markedly more efficient than that observed in the control condition for second-, fourth-, and sixth-grade children.

Two kinds of generality have already been shown for the sentence effect, that is, generality across age and populations. Rohwer (1967, Experiment XIII) has reported significant positive effects of sentence instructions for kindergarten, first-, and third-grade children, both in the case of white subjects drawn from high-socioeconomic status (high-SES) residential areas and for black children drawn from low-SES areas.

Rohwer (1970) has also proposed a formulation giving an hypothetical account of the facilitating effect of sentence instructions. The notion advanced is that the act of constructing sentences for each noun pair prompts internal mental activity that results in modifications of the meaning of the two nouns in a given pair. This modification is thought to consist of investing at least one aspect of each noun with an identical semantic component. Such internal mental activities are referred to with the term elaboration. Thus, sentence instructions are conceived as effective external prompts to internal elaborative activity.

From this point of view, the present experiment was designed to determine whether an analogous elaborative prompt, namely, instructions to create stories, produces facilitation in a different task: the task of free-recall learning. With regard to subject variables, sampling was confined to a single grade level, kindergarten, but did extend across populations, including both high-SES white and low-SES black children.

---

[1] We are indebted to the children, teachers, and administrators in the school districts where the data were collected for their cooperation, tolerance and indispensable help in conducting the study.

Although elaborative prompts were not manipulated, two previous studies have appraised performance on free recall learning tasks among young children sampled from the two populations of interest here. Glasman (1968) administered a list of nouns, represented by the objects denoted by each noun, to samples of high-SES white and low-SES black kindergarten children. The objects were drawn from categories (e.g., clothing, such as coat, dress, sweater; or foods, such as apple, bread meat) but presented in a random sequence during study trials. Performance on this task, as indexed by the number of list items correctly recalled did not vary as a function of population membership. Similarly, in terms of the sequence in which items were recalled a measure of category clustering (the tendency for sequentially adjacent responses to be drawn from the same category) did not vary significantly as a function of population membership.

In a related study (Jensen & Frederiksen, 1970) comparable results were observed when a categorized list of pictured objects (noun referents) was presented in random order to samples of high-SES white and low-SES black second-grade children. Performance was equivalent across the two samples both with respect to the variable of number of items recalled and for that of category clustering. Jensen and Frederiksen also included another version of the same list; the difference was that the items were presented in a blocked order, that is, where all of the members of one category were presented adjacently in the sequence. The results for this categorized-blocked list showed no significant difference between the two samples even though both samples performed at higher levels, as measured by number of correct responses and amount of category clustering, than when the categorized list was presented in random order.

Drawing on these two previous investigations, the present study was planned to assess the effect of elaborative prompts on free recall learning. The lists of items were similar to those used by Jensen and Frederiksen (1970) (categorized-blocked vs. categorized-random) and the samples were similar to those used by Glasman (1968) (high-SES white and low-SES black kindergarten children). Thus, the design permitted an evaluation of the generality of the effect of elaborative prompts both with respect to populations and with respect to list type.

Method

Subjects. The entire sample consisted of 120 children enrolled in kindergarten classes in two schools. One school serves a high-SES white residential area, as defined by census information reported in the 1960 survey, and the other serves a low-SES black residential area. Among the 60 children sampled from each population, the distributions by sex were approximately, although not exactly, equal.

Design. The design was comprised of three factors, each having two levels: Population (high-SES white vs. low-SES black); List (blocked vs. random); and Treatment (story instructions vs. control). Both lists consisted of the same pictures of objects denoted by high-frequency nouns drawn from common semantic categories; the lists differed only with respect to the order in which the items were presented during study trials. The treatment conditions were distinguished by the character of the activity subjects were directed to pursue during specified study-trial pauses: those in the story condition were asked to construct and utter narratives including the nouns to be learned whereas those in the control condition were simply asked to learn the nouns. Within each population, subjects were randomly assigned in equal numbers to each cell of the design (n=15).

Materials and Procedure. The lists consisted of 25 familiar nouns selected from five categories: foods, parts of the body, kitchen utensils, animals, and vehicles. Line drawings were made of the objects denoted by each noun and the drawings were photographed so as to produce 2x2 black-on-white slides. Slide projection was accomplished by means of a Kodak Carousel attached to an audio tape recorder that both controlled the advancement rate on the projector and presented the noun labels of each object in synchrony with the appearance of the corresponding pictures on the screen.

For the purpose of effecting the story manipulation, the lists were subdivided into five parts, quintuples, each consisting of five items. In the blocked list, each quintuple included all of the items from one of the categories; in the random list, each quintuple consisted of one item from every one of the five categories. In all conditions, the presentation sequence during study trials was as follows: the items in the first quintuple were shown at a 4-sec rate, followed by a 15-sec interval, followed by the presentation of the items in the next quintuple, again followed by a 15-sec interval, and so on until all items had been shown and every quintuple had been succeeded by a 15-sec interval. During the 15-sec intervals interposed between quintuples, subjects in the story treatment condition were to construct and utter a narrative of the kind described. Subjects in the control condition were not given specific instructions about what to do during the intervals.

Testing consisted of 3-min, unpaced trials during which subjects were asked to recall and utter all of the nouns presented during the study trial, regardless of order. If subjects paused for longer than 30 seconds and declared their inability to recall more items, test trials were terminated short of the 3-min limit.

A total of three study and three test trials were given. From study trial to study trial, the presentation order for the five quintuples was constant and the membership of each quintuple was the same; the order of items within quintuples, however, was randomly varied across every trial.

All subjects were tested individually. The instructions described the procedure fully and directed the subject's effort to memorizing the items in the list. These instructions were augmented by the presentation of a five-item practice list by means of slides and audiotape. Subjects in the control condition observed the experimenter repeat the five nouns and subjects in the story condition observed the experimenter utter a narrative containing the five nouns (e.g., "The HORSE sat in a CHAIR with his CUP in one HAND and a CAKE in the other."). Those in the story condition were asked to construct and utter similar stories during the intervals between quintuples.

Results

Two measures were used to index performance on the task: the number of items correctly recalled on each test trial; and, the amount of clustering on each test trial. In both cases, two multivariate analyses of variance were performed in which the three dependent variables were the scores on the test trials. The design for the first analysis was chosen to provide maximum sensitivity to treatment effects while the other was optimal for assessing effects associated with the factor of Population. Thus, in the first analysis to be described for each measure, Population was treated as nested within Lists and Treatments were nested in Populations and Lists; in the second analysis, Population was nested in Lists and Treatments. Since the use of two analysis designs inflates the overall Type I error rate, this risk was minimized by testing the List effect only once, by testing the Treatment effect only in the first analysis and by testing the Population effect only in the second.

Items Recalled. The results of the experiment as indexed by the mean number of items correctly recalled on each trial are presented in Table 5. Unexpectedly, the multivariate test for the main effect of

- - - - - - - - - - - - - -

Insert Table 5 about here

- - - - - - - - - - - - - -

lists was not significant ($F = 2.30$, $df = 3/110$, $p > .05$). The effect of story instructions in the blocked list, however, was significant for the low-SES black children ($F = 4.53$, $df = 3/110$, $p < .05$) but not for the high-SES white children ($F = 2.66$, $df = 3/110$, $p > .05$). Even so, univariate analysis of this effect separately for each trial, adopting a stringent error rate ($p = .01$), revealed a significant difference favoring the story condition on trial 3 ($F = 7.44$, $df = 1/112$, $p < .01$). In contrast, the effect of story instructions was not significant on the random list for either low-SES black subjects (multivariate $F < 1$) or for high-SES white subjects (multivariate $F = 1.47$, $df = 3/110$, $p > .05$). Furthermore, univariate analyses for each trial separately did not reveal a significant story effect in any instance for the random list. In summary, the effect of story instructions emerged only on the blocked list and it emerged more generally across trials for the low-SES black than for the high-SES white samples.

The second multivariate analysis design applied to the data revealed only one significant effect associated with the Population factor, namely, that nested in the story treatment of the blocked list. In this condition, the low-SES black children made more correct responses than the high-SES white children ($F = 2.76$, $df = 3/110$, $p < .05$). This outcome, of course, coincides with that showing a positive effect of story instructions for the low-SES black children but not for the high-SES white children on the blocked list. The Population effect was not

Table 5

Mean Number of Correct Response per Trial (max. = 25) as a function of List, Population and Treatment

		High-SES White		Low-SES Black	
List	Trial	Story	Control	Story	Control
	1	9.27	7.73	9.20	7.13
Blocked	2	11.40	8.80	13.27	9.73
	3	13.47	9.73	12.67	11.67
	1	8.53	7.47	7.27	6.40
Random	2	9.40	8.73	10.60	8.80
	3	11.80	9.33	11.47	10.20

significant in any of the remaining three conditions: Control-Blocked ($F = 1.12$, $df = 3/110$, $p > .05$); Story-Random ($F = 2.11$, $df = 3/110$, $p > .05$); and Control-Random ($F < 1$).

Clustering. The amount of clustering in the recall performance of every subject was estimated by computing the Z index proposed by Frankel and Cole (in press). This clustering index is defined as follows:

$$Z = \frac{EM_r - O_r}{\sqrt{EV_r}} ,$$

where r refers to *runs*, that is, the number of times in the sequence of items recalled by a subject that adjacent items were members of different categories plus the constant, 1. Then, O_r is the observed number of runs, EM_r is the expected mean number of runs, and EV_r is the expected variance of runs. The expected values, EM_r and EV_r are computed for all possible sequences of the length observed in recall, given the number of categories represented in the list presented. Thus, Z is a standard score that can be referred to the normal distribution to obtain its probability of occurrence. Clustering is then inferred for these sets of recall sequences that contain significantly too few runs; the smaller the number of runs, relative to the expected number, the higher the Z index.

In the present study, it was of interest to measure the amount of clustering with respect to two kinds of item groups: category groups and story groups. In the case of the blocked list, these two groups were entirely coincident so that the degree of clustering observed may be attributed either to shared membership in a category, to shared membership in a quintuple, or both. In the case of the random list, however, it is possible to separate these two sources of clustering, computing Z separately for category membership and for quintuple, i.e., story membership. The results are presented in Table 6.

- - - - - - - - - - - - - -

Insert Table 6 about here

- - - - - - - - - - - - - -

An application of the first multivariate analysis design to the estimates of category and story clustering provides a sensitive assessment of the effects of story instructions. For the low-SES black children, the simple main effect of Treatment was not significant on the blocked list ($F = 1.78$, $df = 3/110$, $p > .05$). Similarly, among these children, the Treatment effect on the random list was not significant when scored for category clustering ($F < 1$) nor when scored for story clustering ($F < 1$). In the high-SES white samples, however, the Treatment effect was significant on the blocked list ($F = 2.94$, $df = 3/110$, $p < .05$) such that the story condition was associated with more clustering than the control condition. Nevertheless, the Treatment effect was not significant for these subjects on the random list for either the amount of category clustering ($F = 1.85$, $df = 3/110$, $p > .05$) or for the amount of story clustering ($F < 1$). The main effect of

Table 6

Mean Clustering Score (Z) per Trial as a Function

of List, Population and Treatment

		Category Clustering			
		High-SES White		Low-SES Black	
List	Trial	Story	Control	Story	Control
	1	1.80	1.52	1.79	1.12
Blocked	2	2.42	1.14	2.59	1.63
	3	3.10	1.89	2.03	1.43
	1	0.34	0.68	0.56	0.90
Random	2	0.28	0.74	1.00	0.74
	3	1.57	0.68	0.88	0.31

		Story Clustering			
		High-SES White		Low-SES Black	
List	Trial	Story	Control	Story	Control
	1	1.80	1.52	1.79	1.12
Blocked	2	2.42	1.14	2.59	1.63
	3	3.10	1.89	2.03	1.43
	1	0.21	0.20	-0.42	-0.01
Random	2	0.11	-0.25	-0.38	-0.25
	3	0.12	0.07	-0.65	-0.44

List, for the dependent variable of category clustering was significant ($F = 14.41$, $df = 3/110$, $p < .05$) as it was for story clustering ($F = 49.13$, $df = 3/110$, $p < .05$); both variables revealed more clustering on the blocked than on the random list. Even so, the failure of story instructions to produce significant effects on story clustering in the random list contradicts the conclusion that the story manipulation resulted in organization of the list in terms of stories.

The second multivariate analysis design permits an assessment of clustering effects associated with Population. As measured by the index of category clustering, the Population effect was not significant in any condition: blocked-story ($F = 1.47$, $df = 3/110$, $p > .05$); blocked-control ($F < 1$); random-story ($F = 1.87$, df $3/110$, $p > .05$); random-control ($F < 1$). Similarly, none of the tests for Population effects measured by the variable of story clustering was significant: blocked-story ($F = 1.96$, $df = 3/110$, $p > .05$); blocked-control ($F = 1.16$, $df = 3/110$, $p > .05$); random-story ($F = 1.25$, $df = 3/110$, $p > .05$); random-control ($F < 1$).

Discussion

The results of the present experiment suggest that a technique of elaborative prompting, i.e., instructions to generate stories, has generality to the task of free recall learning. This generalization, however, is narrowly limited to the blocked method of presenting a categorized free recall list, at least among kindergarten children. As indexed by the number of correct responses made, the effect on the blocked list held for both populations sampled, low-SES black and high-SES white, although, for the latter sample the effect was evident only on trial 3. In contrast, the sequence of item recall reflected greater category clustering in the story condition only for the high-SES white children.

One interpretation of the effect of story instructions on learning the blocked list is that subjects were prompted to organize and retrieve the list items in terms of the stories they had produced; the category structure of the list is, thus, relegated to an ineffectual status. This interpretation, however, implies that story instructions should have been as effective for learning the random list as they were for learning the blocked list. The results disconfirm this implication. Furthermore, the index of story clustering revealed no effect of story instructions in the case of the random list--apparently, subjects did not organize and retrieve items in connection with their narratives. An alternative interpretation is more viable, namely, that the story instructions served to make more salient the category membership of the items in the blocked list, prompting the subjects to arrange the items in terms of categories for the purpose of increasing their retrievability. This interpretation clearly is conjectural and presently awaits experimental verification.

The limited generality of the elaborative prompting technique manipulated in the present study appears confined to the young age level sampled here. In a closely related experiment, the effects of story instructions among samples of high-SES white fifth-grade children was to facilitate learning on both a random and a blocked list as well as on a list of unrelated items (Irwin, Gerdes & Rohwer, 1971). Thus, a permissible inference from the present results accords well with one made from the results of a study designed to train preschool and kindergarten children in the use of elaborative prompts for noun-pair learning: it may be more productive to delay training in the autonomous use of elaborative prompts until pre-adolescence and to insure efficient learning at younger age levels by incorporating the prompts directly in the materials presented for learning.

With regard to the matter of differences between populations in the efficiency of free recall learning, the present results are consistent with those of experiments reported previously (Glasman, 1968; Jensen & Frederiksen, 1970). At ages as young as kindergarten, there are virtually no discrepancies between high-SES white and low-SES black children in their proficiency at free recall learning. In the present experiment, it is notable that the only significant Population effect, that in the blocked-story condition, favored the low-SES black over the high-SES white children. It remains a mystery that population differences in performance emerge at the kindergarten level for the task of noun-pair learning but not for the task of learning a list of nouns by the method of free recall.

Individual Differences in the Learning of Verbally and Pictorially Presented Paired Associates[1,2]

Joel R. Levin, William D. Rohwer, Jr. and T. Anne Cleary

Current educational technology makes it possible to provide for individual differences in learning as never before. Since the introduction of branching programs for simple teaching machines (Crowder, 1960), individualized speed-sequence-content information (based on student entering behaviors and ongoing performance) is now easily incorporated into classroom instruction through the continual data processing/program modification of computerized systems. (See, for example, Suppes and Morningstar, 1969).

However, despite well-founded pleas that individual abilities be given greater regard in the context of human learning (Cronbach, 1967; Cronbach & Snow, 1969; Jensen, 1967; Messick, 1969), few efforts in the laboratories have been directed toward this end. While a host of experimental variables are known to affect performance on particular learning tasks, typically these have been discovered by comparing the average scores of two or more independent groups. In such experiments, the variation attributable to subjects within groups constitutes the error term and, if substantially smaller than the between group variability, produces a significant treatment effect.

While such outcomes may provide information of some general utility, one should be reminded by Jensen's (1967) comment:

> "Only if it has been demonstrated that the Subjects X Independent Variable interaction is negligible can we be very sanguine about the psychological importance of a particular independent variable, when our conclusions are based on group mean differences. It is preferable to know what happens to individuals under the effect of the independent variable. Experimental psychologists are not interested fundamentally in group effects. Our aim essentially is to devise experiments that will yield information capable of narrowing the range of alternative models of the mind."

Investigators sometimes remark that an error term seems "excessively large" or "remarkably small," but seldom is this source of variation studied systematically; that is, in a way which would suggest the extent to which individuals may be reliably classified according to their

---

[1] This work was supported, in part, by a grant from the National Institutes of Health (HD03869-01). The paper was prepared while Dr. Levin was affiliated with the Wisconsin Research and Development Center for Cognitive Learning. The authors are grateful to Miss Carole King and Miss Caryl Sutton who helped compile the data, as well as to Dr. Herbert J. Walberg for his suggestions based on an earlier version of this paper.

[2] Originally published in American Educational Research Journal, 1971, 8, 11-26.

relative performance under different treatment conditions. Although the usual independent groups experiment does not permit this kind of inference, designs and methods appropriate for determining whether variation attributable to individual differences is reliable do exist.

Investigations of verbal learning in which properties of the stimulus materials are varied have made significant contributions to the study of group effects, but surprisingly little to the study of individual effects. One domain which has received considerable attention, of late, concerns the comparative efficacy of study materials which are either verbal (presented aurally or in printed form) or pictorial (presented visually), (Davidson & Adams, 1970; Dilley & Paivio, 1968; Fredrick, Blount & Johnson, 1968; King, Roberts & Kropp, 1969; Milgram, 1967b; Milgram & Riedel, 1969; Paivio & Yarmey, 1966; Rohwer, Lynch, Levin & Suzuki, 1967; Rohwer, Lynch, Suzuki & Levin, 1967; Stevenson & Siegel, 1969; Yuille & Paivio, 1968). In such studies, concluding statements like "Pictured objects produced more efficient learning than did labeled objects", "A symbolic description of set theory was more effective than was a figural description", or "Sentence and motion picture presentations were equally efficacious" provide a helpful starting point. At the same time, however, a given treatment is unlikely to be uniformly beneficial for every S to whom it is administered. For example, verbal and pictorial materials may produce different patterns for relative benefit for Ss of varying ages, sexes, IQs and social classes. In fact, classification of Ss along such dimensions is typical of this kind of research.

The quest for aptitude by treatment interactions (ATI), where different intellectual and conceptual abilities are hypothesized to interact with various instructional methods has become increasingly popular in recent years. In terms of the verbal-visual dichotomy, one might expect that Ss with particular aptitudes or "preferences" for stimulus inputs in one sensory modality (e.g., aural as opposed to visual) will have a greater probability of succeeding in specific learning tasks if materials are presented in a manner congruent with their preferences. Unfortunately, aptitude-treatment interactions are not easily demonstrated (cf. Cronbach & Snow, 1969).

One problem with the studies conducted thus far is that the choice of aptitude to define, let us say, "verbal" or "pictorial" preferences is frequently only superficially related to the criterion behavior of interest. Thus, it is not unusual to use Ss' performance on a paper folding task (perhaps labeled "spatial ability") to predict the learning of basic algebra with or without benefit of graphs (Carry, 1967). Cronbach & Snow (1969), after receiving the ATI literature accumulated to date, hasten to point out that "...simple characterizations of aptitudes and treatments in such terms as 'spatial' are unlikely to identify combinations of variables worth investigating."

A major contention of the present authors is that before pursuing ATI possibilities within a given area, one must first demonstrate the reality of the phenomenon in its simplest form: that is, the aptitude selected should be defined by performance on a task which is identical (or very similar) to that being predicted. In other words, if one is interested in predicting learning, an index based on how well a child can learn (i.e., performance on a learning task) would be more suitable than one based on an IQ or aptitude test which is generally regarded as revealing what he has already learned (Jensen, 1969; Rohwer, 1970).

The results to be presented here consist of a subsidiary analysis of data collected from a large-scale study reported by Rohwer, Ammon, Suzuki & Levin (1971), where a more complete description of the materials and procedures may be found.

In the Rohwer et al.(1971) paper, only group differences in stimulus preference were investigated; whereas in this paper, individual (within group) differences will be considered. Of primary concern is the consistency (over time) with which individual Ss prefer or benefit more from learning materials presented in one form as opposed to another (in this case, verbally or pictorially).

Method

The Learning Task. In order that individual stimulus preferences in learning could be examined, a paired-associate (PA) task was used in which the stimulus materials consisted of pairs of familiar objects. All materials were filmed and subsequently displayed (with audio) on a videotape monitor.

Five different types of items were prepared, the basic distinction among the types being the amount of verbal and/or pictorial support provided for the pairs. Specifically, for two item types, the paired objects were either named aloud on the tape in the absence of pictorial support [Verbal (V)] or pictured adjacently on the monitor in the absence of verbal support [Pictorial (P)]. The three other item types consisted of combinations and elaborations of these verbal-pictorial descriptions: named on the tape while pictured adjacently on the monitor; related to one another verbally—via a sentence on the tape—while pictured adjacently on the monitor; or named on the tape while related to one another pictorially—via an animated sequence on the monitor.

Two versions of the PA task were created by randomly assigning items to two forms (A and B), with each form comprising two lists of 25 PAs, five of each item type. A random sequence of the 25 PAs within each list was constructed after having randomly determined which pairs would be presented as which item types.

Procedure. All Ss were tested individually, and were provided with two alternating study and test trials for each list (with items appearing in different random orders on each trial, to control for serial learning). Items were presented at a 4-second rate, with an 8-second interval between study and test trials. Form A of the test (Lists 1 and 2) was administered during one session, with Form B (Lists 3 and 4) administered two days later, in order that the stability of PA learning proficiency (and preferences) could be assessed.

Subjects. A total of 288 Ss was obtained for the present study, with 24 males and 24 females randomly selected from three grades (kindergarten, first grade, and third grade) in two different schools; one of the schools served a relatively high socioeconomic status (SES) white community, and the other a low-SES black community.[3]

Results

Identification of Verbal and Pictorial Preferences. The title of this paper describes a particular kind of item-type preference that is of interest here: viz., verbal vs. pictorial. For this reason, only two of the five item types will be considered further, V items (unseen pairs that were labeled) and P items (pictured pairs that were not labeled). Each of the other item types incorporated both verbal and pictorial components in varying degrees, while the two retained included only one of these components. Thus, by excluding items consisting of auditory and visual combinations, our discussion will focus on the two "pure" versions of verbally and pictorially presented materials.[4]

One way in which item-type preferences may be examined is to group individuals on the basis of their V and P item-type performance for Form A of the PA task ("classifying" lists), then determine whether such groupings are effective predictors of item-type performance on Form B ("criterion" lists). Operationally, a "preference" has been defined in terms of S's relative performance on P and V items; that is, the difference between his recall of P and V items (P-V).

Classifications of High and Low P-V Types were made within SES-race, age, and sex groups. A total of 12 (two levels of SES-race, three of age, and two of sex) reference samples were therefore derived, with each initially containing 24 Ss.

High P-V Types included all Ss for whom the difference, P-V, was above the median of their reference sample, while Low P-V Types consisted of Ss whose P-V differences were below the median. (Ss whose difference scores fell at the median were discarded.)

The performance of High and Low P-V Types on the classifying lists (Form A) is found in Table 7, broken down by item types (V and P) and

- - - - - - - - - - - - - -

Insert Table 7 about here

- - - - - - - - - - - - - -

reference groups. It may be noted that P-items were, in general, easier to learn than V-items, so that the distribution of P-V differences does not center about zero. Thus, Low P-V Types typically consisted of Ss for whom the inequality P > V was relatively small, rather than reversed as in V > P. At the same time, Table 7 reveals that while large P-V

---

[3] It should be kept in mind that the two factors "SES" and "race" are confounded, which is true of most research of this kind. The compound label, "SES-race," will therefore be used to remind the reader of this.

[4] This is not meant to imply that supplementary elaborative strategies were not employed by S in learning the pairs by engaging other verbal or imaginal processes during storage. Rather, "pure" refers explicitly to the particular manner (verbal or pictorial) in which the pairs were presented.

Table 7

Classifying List Performance of Hi and Lo P-V Types on V and P Items (Sample sizes in parentheses)

		Learner Type	*V Items*	*P Items*	*Sum (P+V)*	*Difference (P−V)*
Kindergarten	**High-SES White:**					
	Boys	Hi P-V (12)	1.83	6.42	8.25	4.59
		Lo P-V (12)	4.17	3.42	7.59	−0.75
	Girls	Hi P-V (11)	1.00	5.09	6.09	4.09
		Lo P-V (9)	1.78	1.11	2.89	−0.67
	Low-SES Black:					
	Boys	Hi P-V (11)	0.73	5.64	6.37	4.91
		Lo P-V (8)	1.25	0.88	2.13	−0.37
	Girls	Hi P-V (10)	1.00	4.80	5.80	3.80
		Lo P-V (11)	0.82	0.27	1.09	−0.55
1st Grade	**High-SES White:**					
	Boys	Hi P-V (11)	1.73	7.54	9.27	5.81
		Lo P-V (11)	2.91	3.45	6.36	0.54
	Girls	Hi P-V (10)	1.30	6.50	7.80	5.20
		Lo P-V (7)	2.86	2.71	5.57	−0.15
	Low-SES Black:					
	Boys	Hi P-V (10)	1.50	8.00	9.50	6.50
		Lo P-V (8)	1.50	1.50	3.00	0.00
	Girls	Hi P-V (12)	1.17	5.92	7.09	4.75
		Lo P-V (12)	1.50	1.67	3.17	0.17
3rd Grade	**High-SES White:**					
	Boys	Hi P-V (11)	3.00	9.73	12.73	6.73
		Lo P-V (11)	6.18	6.45	12.63	0.27
	Girls	Hi P-V (11)	2.82	8.64	11.46	5.82
		Lo P-V (8)	7.00	6.50	13.50	−0.50
	Low-SES Black:					
	Boys	Hi P-V (11)	1.09	9.09	10.18	8.00
		Lo P-V (9)	3.56	4.67	8.23	1.11
	Girls	Hi P-V (10)	1.10	6.00	7.10	4.90
		Lo P-V (11)	4.09	3.36	7.45	−0.73

differences tend to accompany superior total performance (the sum of V and P items, or P + V) among younger Ss (kindergarten and first grade), this is not the case among older ones (third grade). While younger High and Low P-V Types appear to differ in both discrepancy between and performance level on V and P items, the latter difference is negligible among older Ss.

Criterion Performance of the Preference Groups. In Table 8 is presented the performance on the criterion lists (Form B) of the initially

Insert Table 8 about here

classified High and Low P-V types in each of the reference samples.

The two variables, P and V, were transformed into two new measures for analysis-of-variance purposes: one, the sum of the two variables (P+V), was used to test for Between-Learner-Type differences; while the other, the difference between the two (P-V), was used to examine Learner Type X Item Type interactions. If individual preferences are indeed stable, the latter test would be expected to detect a significant interaction, since Ss were assigned to High and Low P-V groups on the basis of the size of P-V differences on the classifying lists (Form A).

The data were analyzed as a nested design, with differences between High and Low Types tested within each reference sample. All tests of hypothesis were performed with a specified Type I error rate of .05.

Concerning performance level (P+V), in only one of the twelve reference samples (Low-SES Black Kindergarten Boys) was a statistically significant difference Between-Learner-Types detected ($F = 5.24$ with 1 and 223 d/f, $p < .05$). From Table 8 it may be observed that this represents a difference of almost five items (7.09 correct for High P-V Types versus 2.38 correct for Low P-V Types).

More germane to the present study were the significant Learner Types X Item Types interactions that were obtained in five of the twelve reference samples. These appeared among High-SES White Kindergarten Boys ($F = 7.20$ with 1 and 223 d/f, $p < .01$), among High-SES White First Grade Boys ($F = 10.46$ with 1 and 223 d/f, $p < .005$), and among each of the Low-SES Black male samples: Kindergarten ($F = 8.19$ with 1 and 223 d/f, $p < .005$); 1st Grade ($F = 6.31$ with 1 and 223 d/f, $p < .05$); and 3rd Grade ($F = 17.32$ with 1 and 223 d/f, $p < .001$).[5] In sum, in 10 of the 12 reference groups, the differences were in the predicted direction.

Although the classification into Learner Types was based on relative preferences, it is interesting to note that in four of the five samples with significant Learner-Types X Item Types interactions, Low P-V Types did retain (though not assessed statistically) their absolute superiority on V items. Also, even though significant interactions were not demonstrated in all samples, an inspection of Table 8 reveals that in nine out of twelve cases Low P-V Types recalled more V items than did High P-V Types; and in another nine out of twelve cases, High P-V Types

Table 8

Criterion List Performance of Hi and Lo P-V Types on V and P Items (Sample sizes in parentheses)

		Learner Type	*V Items*	*P Items*	*Sum (P+V)*	*Difference (P−V)*
Kindergarten	**High-SES White:**					
	Boys	Hi P-V (12)	2.50	5.42	7.92	2.92*
		Lo P-V (12)	3.83	3.17	7.00	−0.67
	Girls	Hi P-V (11)	1.54	3.45	4.99	1.91
		Lo P-V (9)	1.56	4.44	6.00	2.88
	Low-SES Black:					
	Boys	Hi P-V (11)	1.18	5.91	7.09	4.73*
		Lo P-V (8)	1.00	1.38	2.38*	0.38
	Girls	Hi P-V (10)	0.70	3.30	4.00	2.60
		Lo P-V (11)	1.36	3.73	5.09	2.37
1st Grade	**High-SES White:**					
	Boys	Hi P-V (11)	2.64	7.73	10.37	5.09
		Lo P-V (11)	2.91	5.91	8.82	3.00
	Girls	Hi P-V (10)	2.00	6.50	8.50	4.50*
		Lo P-V (7)	3.00	2.28	5.28	−0.72
	Low-SES Black:					
	Boys	Hi P-V (10)	1.50	6.90	8.40	5.40*
		Lo P-V (8)	2.62	4.12	6.74	1.50
	Girls	Hi P-V (12)	1.83	4.75	6.58	2.92
		Lo P-V (12)	2.00	2.83	4.83	0.83
3rd Grade	**High-SES White:**					
	Boys	Hi P-V (11)	4.27	9.18	13.45	4.91
		Lo P-V (11)	3.82	7.36	11.18	3.54
	Girls	Hi P-V (11)	4.73	7.36	12.09	2.63
		Lo P-V (8)	4.62	7.62	12.24	3.00
	Low-SES Black:					
	Boys	Hi P-V (11)	2.82	9.27	12.09	6.45*
		Lo P-V (9)	4.89	5.22	10.11	0.33
	Girls	Hi P-V (10)	1.90	6.80	8.70	4.90
		Lo P-V (11)	3.00	5.18	8.18	2.18

* Significant with $\alpha = .05$

recalled more P items than did Low P-V Types.

The method of classification used here, that is, assigning Ss to High and Low P-V groups is only one system which might have been employed. It was originally hoped that approximately equal numbers of "V-Preference" and "P-Preference" Ss could be identified, but this possibility was ruled out by the fact that P items were much easier to learn than were V items. An inspection of Form A performance, presented in Table 9, reveals that in comparison to P items, V items tended to produce a reduced spread of scores, in some cases (viz., the younger low-SES black samples) resulting in a marked floor effect. (Since these figures are based on five item types summed over two lists and two trials, the maximum score possible is 20.) Thus, if classification of Ss on the basis of absolute preferences had been adopted, extremely few "V preference" Ss would have been located.

- - - - - - - - - - - - - -

Insert Table 9 about here

- - - - - - - - - - - - - -

For this reason, the relative preference scheme was employed. One difficulty with this system is that while High P-V differences may accurately reveal the preference variable being considered, Low P-V differences identify individuals whose performance was (a) high on both P and V items, (b) low on both P and V items, or (c) somewhere between these extremes. As a result, it is likely that group preferences would be confounded with the levels at which Ss in the two groups were performing.

Certainly other classification procedures were possible. For example, one might assign Ss to one of four mutually exclusively categories, according to their initial joint P and V item performance: High P, High V; High P, Low V; Low P, High V; and Low P, Low V. However, as Cronbach (1968) has recently argued in reference to the Wallach and Kogan (1965) data, categorization procedures of this kind involve discarding potentially relevant information in the data, thereby leading to unparsimonious and/or inappropriate interpretations.

For these reasons, the present data will now be analyzed correlationally, retaining the original score information.

Reliability of Preferences. In the pages that follow, inter- and intra-form correlations between V and P items will be examined. The "sex" factor has been excluded from the remaining analyses, since it was found not to interact with the variable of interest, namely item-type preferences.

With two item types (V and P) and two forms (A and B) of the PA

---

[5] Since the nature of the Learner Type X Item Type interaction is clearly directional, if a one-tailed test of significance had been employed, a significant effect would also be detected among low-SES Black 3rd Grade Girls ($F = 3.61$ with 1 and 223 d/f, $p < .05$).

Table 9

Means and Standard Deviations of Form A Verbal (V) and Pictorial (P) Items for the Six Samples (48 Ss per sample)

	Item Type	*Mean*	*S.D.*
Kindergarten:			
High-SES White	V	2.21	2.06
	P	4.06	2.74
Low-SES Black	V	0.81	1.30
	P	2.71	2.84
1st Grade:			
High-SES White	V	2.08	1.90
	P	5.02	2.79
Low-SES Black	V	1.40	1.67
	P	4.35	3.26
3rd Grade:			
High-SES White	V	4.25	3.18
	P	7.54	3.29
Low-SES Black	V	2.50	2.57
	P	5.92	2.94

task, six different zero-order correlations were computed:

(a) Two, described as item-type reliabilities, are the correlations between scores for a particular item type on Form A and scores for the same item type on Form B which was administered two days later (V_AV_B and P_AP_B).

(b) Another two, labeled inter-form correlations, are the correlations between scores on a particular item type on Form A and scores on the other (different) item type on Form B (V_AP_B and P_AV_B).

(c) A final two, called intra-form correlations, are the correlations between scores on the two different item types within each form of the test (V_AP_A and V_BP_B).

The correlation coefficients for each of the reference groups are presented in Table 10, along with the averages of the two coefficients

- - - - - - - - - - - - - - -

Insert Table 10 about here

- - - - - - - - - - - - - - -

within each classification.[6] Descriptively, the correlational patterns for each sample in Table 10 reveal item-type reliabilities larger in magnitude than either the inter- or intra-form correlations. The fact that within item-type correlations over time (item-type reliabilities) exceed the between item-type correlations at the same point in time (intra-form correlations) suggests that stable verbal-pictorial preferences exist among these children. No evidence to the contrary is found in any of the six samples investigated.

Finally, the relationship between Form A and Form B item-type preferences may be defined in terms of the relative performance exhibited by Ss on each occasion. Paralleling the P-V subject classification employed earlier, Form A difference scores were computed for each S by subtracting his V score from his P score. This was done in two ways: with a simple raw-score difference and with a difference obtained after equating the within-group standard deviations of the item-type scores. When Form A and Form B difference scores were correlated, an index of "relative preference" stability was determined.

The reliabilities of the difference scores, both raw and standardized are presented in Table 11. Given that the reliability of difference

- - - - - - - - - - - - - - -

Insert Table 11 about here

- - - - - - - - - - - - - - -

scores is usually quite low, these coefficients are of reasonable magnitudes, especially in the Low-SES black samples. Furthermore, the

---

[6] Simple means of the r's are reported here, since they were found to be almost identical to the averages obtained when Fisher's Z-transformation was employed.

Table 10

Verbal and Pictorial Item-Type Correlations for the Six Samples

	Item-Type Reliabilities (Same item type, different time)			*Inter-form Correlations (Different item type, different time)*			*Intra-form Correlations (Different item type, same time)*		
	V_AV_B	P_AP_B	*Mean*	V_AP_B	P_AV_B	*Mean*	V_AP_A	V_BP_B	*Mean*
Kindergarten:									
High-SES White	.45	.24	.34	.07	.17	.12	.26	.09	.18
Low-SES Black	.59	.56	.58	.31	.20	.26	.33	.50	.42
1st Grade:									
High-SES White	.45	.56	.50	.30	.19	.25	.36	.02	.19
Low-SES Black	.38	.50	.44	.26	—.03	.12	.36	.22	.29
3rd Grade:									
High-SES White	.50	.51	.50	.16	.53	.34	.40	.47	.44
Low-SES Black	.62	.60	.61	.12	.11	.12	—.01	.26	.12

Table 11

The Reliability of Raw and Standardized Differences between P and V Items

	Correlation Between	
	Raw P-V Differences	*Standardized P-V Differences*
Kindergarten:		
High-SES White	.24	.27
Low-SES Black	.53	.53
1st Grade:		
High-SES White	.34	.32
Low-SES Black	.51	.46
3rd Grade:		
High-SES White	.27	.28
Low-SES Black	.56	.56

reliabilities of the raw and standardized measures are quite comparable within each sample.

Discussion

The Future of Verbal and Pictorial Preferences. The data considered herein are suggestive of individual mode (in this case, verbal-pictorial) preferences, at least with respect to the present task and populations. Even though the recall of unlabeled pictorial PAs was generally greater than that of unseen pairs which were named, substantial individual difference variation was observed in each of the six age/SES-race samples. For example, there were several children who exhibited a "verbal over pictorial" advantage which held up on both forms of the test.

The fact that different stability indices emerged among the various age and SES-race samples requires additional exploration. The possibility of mode preferences interacting with the SES-race variable, for example, should not be easily dismissed.

Further assertions are not warranted, however, until both replications and modifications of the present research are completed. For one, task characteristics should be subjected to examination. In the study reported here, following the study trial Ss were required to vocally supply the missing response object when the appropriate stimulus object was presented (verbally or pictorially). For V items, Ss initially heard the object names; and for P items, they initially saw the object pictures. Yet for both item types, Ss were subsequently asked to utter the missing object's name aloud. Whether this kind of procedure is prejudicial to a particular item type should be verified by means of a systematic investigation of study/test trial modes of presentation and testing. In this regard, picture recognition may well be the most appropriate method for assessing P-item learning.

Whether or not verbal and pictorial preferences extend beyond labeled objects or static representations of them is another interesting question, which was not conducive to investigation here since the item types other than V and P each included stimuli in more than one mode. Some work currently underway, in which items consisting of sentence-embedded PAs without pictorial support and pictorially interacting PAs without verbal support are being used, might provide an answer. Between- and within-subject factorial manipulations of item types (V or P) in conjunction with the presence or absence of provided elaboration in the same (or different) mode might also unearth the kind of spontaneous strategies utilized by children during PA and other types of learning.

Finally, the present research appears promising with respect to its potential for classroom utilization. Diagnosis of individual learner types could be used to determine which kinds of learning materials are most suitable for which kinds of students. Following individual diagnosis, two complementary approaches might be recommended: (a) the selection of learning materials which best match the student's learning preferences, or (b) the amelioration of deficient preferences (or the elicitation of inoperative ones) through training, as related to existing or desired learning materials.

Just this kind of attention to identifying and making provisions for different types of learners is necessary for instruction to be truth "individualized." Until learner types, in addition to cognitive styles and affective reactions (Messick, 1969), are more carefully attended to, school instruction will continue to be less than optimally efficient.

Effects of Training in Vocabulary vs. Sentence Construction[1]

Paul R. Ammon and Mary Sue Ammon

In recent years there has been considerable interest in facilitating the language development of young disadvantaged children through some sort of direct intervention. In spite of much activity in this area, very little is known about the differential effects produced by specific types of early language training. It is important to find out which aspects of language are most amenable to early training, and which types of training have the greatest transfer value for stimulating language development in general. Otherwise, attempts at intervention will be inefficient, or even inappropriate, with potentially damaging consequences for the children involved.

This chapter of the present report is concerned with an experimental study which compared the effects of training young black children in vocabulary versus sentence construction. The rationale for this experiment was based, first of all, on the observation that previous studies have dealt with relatively complex language training programs, making it difficult to infer specific causes of the effects obtained. Experimental studies with a practical orientation have sought to maximize contrasts between experimental and control groups or between experimental treatments representing wholly different approaches to language training. Consequently they have confounded a variety of teaching methods and curricular contents within each treatment (e.g., Klaus & Gray, 1968; Dickie, 1968). Interestingly, these efforts to accentuate the differences between treatments have not always produced notable differences in results (e.g., Dickie, 1968). Even experiments with a more theoretical orientation have confounded the effects of instructional method and content (e.g., Cazden, 1965). The strategy of the present study was to hold method as constant as possible, while setting up a clear-cut difference in content. The specific contrast between vocabulary and syntax was suggested not only by a formal analysis of language, but also by some recent theorizing about language acquisition and the disadvantaged child. On the one hand, vocabulary has been singled out as the major area of language in which young children normally receive much deliberate instruction from adults (Cazden, 1968). Thus, if a child grows up around adults who do not pay much direct attention to his language, vocabulary may be the key to any language deficit which results. On the other hand, one popular belief nowadays is that disadvantaged children are brought up in a community which speaks mainly a "restricted code" (Bernstein, 1964). The restrictedness of the code is largely a matter of syntax, so that the child does not develop an ability to use the full range of syntactic possibilities for "elaborating" his sentences with words, phrases, and clauses which make meaning more explicit.

---

[1] The findings of this study were presented by Ammon and Ammon (1970) at the annual meeting of the American Educational Research Association, March 1970. A paper based on the data from preschool children has been accepted for publication in the Journal of Educational Psychology (Ammon & Ammon, in press). The authors are grateful to Dr. Joel Levin, Dr. Nancy Suzuki, and Sylvia Zuck for their patient help in coding and analyzing the data. Sharon Alexander, Carol Meredith, Sharon Rose, and Caryl Sutton assisted in the testing and training phases of the research. Phyllis Koppelman and Pat Stohr transcribed much of the speech data.

The available data on language abilities among disadvantaged black children are not very helpful in pinpointing the source of the deficiency which these children are reputed to have. They usually attain low scores on standard tests of vocabulary (Cazden, 1966; Raph, 1965), but these scores may simply reflect a symptom rather than a primary cause of deficiency. And although Bernstein's notion of a restricted code has often been used to characterize the language of lower-class black children (Deutsch, 1963; Bereiter & Engelmann, 1966; Osborn, 1968), it is not clear how legitimate this characterization is. There has been a tendency to interpret deviations from standard English as evidence of a restricted code. A distinction must be made between the features of a restricted code (as defined by Bernstein) and the features of a nonstandard dialect spoken in the black community. Labov (1970) and other sociolinguists have argued convincingly that Black English, defined as a system of grammatical rules, is no more restricted or impoverished than Standard English. Furthermore, it can be shown that black speakers produce some very complex speech, beginning at an early age. Linguistic analyses of this sort have led some writers to propose the use of a "difference model" rather than the more usual "deficit model" in comparing lower-class blacks to middle-class whites (e.g., Baratz, 1969). But while a difference model does seem more appropriate for many social-class and ethnic comparisons, the deficit model may still have some validity. Disadvantaged black children may, in fact, be somewhat less skilled than their middle-class peers in using the syntactic and lexical resources of their own communities to express themselves. The evidence of such deficits is rather flimsy -- especially in the area of syntax -- but so is the counter evidence that there are no deficits, only differences.

Given the present uncertainty over the language abilities of disadvantaged black children, a deficit model might at least have some heuristic value. If there is a deficiency, it seems reasonable to suppose that the most effective intervention will be a direct attack on the root cause of the problem. As indicated above, both vocabulary and syntax are conceivable sources of language deficiency in lower-class children. Since it is not clear whether the disadvantaged black child's putative deficiency stems primarily from one or the other (or from both), a study comparing the effects of training in vocabulary versus sentence construction seems worthwhile. The results might not only provide guidelines for early intervention programs, but also shed new light on the basic controversy over differences and deficits in language development.

Method

Subjects.[2] Two large samples of black children participated in two separate but parallel experiments.

---

[2] The authors wish to thank the administrators and teachers who provided the necessary subjects and facilities in the Richmond and Berkeley schools.

The first sample consisted of 72 children, half boys and half girls, who were drawn from seven pre-kindergarten classes at three elementary schools in Richmond, California. At the time of the study, nearly all of the children attending these schools were black, and most of them came from lower-class homes, as indicated by census tract data for the area. The total enrollment in the seven classes was approximately 95. Some of these children were eliminated from the study because they were not black, because they were observed to have severe speech impediments, or because they did not attend school regularly. The final sample of 72 was determined by the feasibility of scheduling training sessions which met the design requirements outlined below. Where more than one child was available for a given position, random selections were made. The selected subjects had a mean age of 4 1/2 years with a standard deviation of 3.3 months when the study began.

The second sample consisted of 72 kindergarten children, half boys and half girls, who were drawn from nine classes at two schools in Berkeley, California. The schools in this district were racially integrated, so that each class contained roughly equal numbers of black and white children. The total enrollment of black children in the nine classes was approximately 87. Most of these children came from lower- to lower-middle-class homes. The final sample of 72 children was selected on the same bases as with the preschool sample. When the study began, the kindergarten subjects had a mean age of 4 years 9 months, with a standard deviation of 3.7 months.

It should be noted that the preschool and kindergarten samples differed in at least three potentially important ways: the kindergarteners were older, had more exposure to white children and their language in school, and were somewhat higher on the usual indices of socioeconomic status.

Overall Design and Procedure. A pretest battery of language instruments was administered individually to each subject. Within each grade level, the children were then divided equally into three treatment groups: vocabulary training, sentence training, and control. The subjects in the two training groups were further divided into groups of four, with two boys and two girls per group when possible. Each of these smaller groups met with the same experimenter twice a week for a 20-minute training session outside of the regular classroom. The training period lasted six weeks, during which time the control subjects had virtually no contact with the experimenters. Finally, all subjects were posttested with the original battery of language instruments.

The experimenters in each phase of the study were three white, college-educated women who had no connection with the regular school program. The subjects in each treatment group were assigned to experimenters A, B, and C according to the design shown in Table 12. (It should be noted

- - - - - - - - - - - - - -

Insert Table 12 about here.

- - - - - - - - - - - - - -

that experimenter C was actually a different person in each phase of the study, due to unforeseen circumstances.) This design balanced the

Table 12

Assignment of Subjects in each Treatment Group

to Experimenters A, B, and C

Phase of study	Treatment Group								
	Vocabulary			Sentence			Control		
Pretest	A	B	C	A	B	C	A	B	C
Training	B or C	A or C	A or B	B or C	A or C	A or B	--	--	--
Posttest	A	B	C	A	B	C	A	B	C

effects of different experimenters on testing and on training, and it permitted a "blind" posttest. As an additional control, each of the regular teachers had approximately equal proportions of children in all three treatment groups. Within all of these constraints, there was random assignment of subjects to testers and to treatment groups.

Materials. Three language instruments were administered as pre- and posttests. The first was Form B of the Peabody Picture Vocabulary Test (PPVT), a widely used recognition test of vocabulary which needs no further description here. The second instrument was a Sentence Imitation Test (SIT) consisting of 50 model sentences which were constructed especially for the present study. In keeping with the focus on syntactic elaboration, these sentences were designed to assess a child's control of several elaborative sentence elements, including prenominal adjectives and possessive nouns, locative prepositional phrases, relative clauses, and adverbial clauses. The experimenter presented each sentence in Standard English with the instruction to "say just what I say." Finally, there was a Picture Interview (PI) patterned after the one used by Loban (1963). In the PI, the experimenter first asked the child some questions about his family, favorite television programs, and so on. Then a series of six pictures was shown and the child was encouraged to talk about each one. All responses to the SIT and to the PI were tape recorded for later transcription and scoring.[3]

The primary material for vocabulary training was a list of target words drawn from two sources (See Appendix B). Thirty-six of the words were taken from the first 60 items in the PPVT-Form B. Previous data had indicated that these words were unknown by a substantial proportion of children similar to those in the present study. Such blatant "teaching to the test" was a way of establishing whether or not the vocabulary training had any effect at all. Another set of 33 words was suggested by the pictures in the PI. That is, these words could be used to describe or discuss the pictures. The primary material for sentence training consisted of the set of elaborative constructions assessed by the SIT (See Appendix C). In teaching these constructions, an attempt was made to use only words which were already familiar to the subjects. A large collection of objects and pictures was developed in order to provide referents for the words and sentence constructions being taught. Each experimenter was equipped with essentially the same set of materials. At no time during training were the test pictures or sentences introduced.

Training Procedures. All three experimenters worked from a common lesson plan for each training session (See Appendix D for sample lesson plans.). The methods used for vocabulary training and for sentence training were quite similar. Imitation and group responding were used somewhat to introduce new forms, but the emphasis was on having individual children respond to a variety of referents for each target form. Both recognition and production tasks were employed in the context of game-like activities, with the ultimate goal being for each child to generalize the production of appropriate words or sentences to new referents. An attempt was made to give every subject an equal number of opportunities to respond. Correct responses were praised and corrective feedback was provided when a form was used incorrectly or when there was no response. Although practically all of the children used some elements of Black English, the sentence training focused not on dialect features, but on the use of elaborative sentence elements which occur in both the standard and the nonstandard dialects of adults. Thus the white experimenters did not emphasize the differences between their own dialects and those of the children.

---

[3] Both the SIT and the PI are discussed in much greater detail in the next chapter of this report.

Analysis. It was assumed that the random assignment of subjects to treatment groups would preclude significant group differences on the pretest, and the pretest data supported this assumption. Consequently, the effects of training were assessed by comparing group means on several dependent variables from the posttest. The data from preschool and kindergarten subjects were analyzed separately. A univariate analysis of variance was done for each dependent measure, with two planned comparisons at the .025 level of significance (Hays, 1963). In these analyses, the N for each group was 23 or 22, due to incomplete data on some subjects.

Results

Preschool Results

Table 13 contains the group means for preschool subjects on five dependent measures.

Insert Table 13 about here.

Effects of Training on Vocabulary Recognition. The PPVT yields IQ and MA scores, but these normative scores were rendered meaningless by the teaching of test words to subjects in the vocabulary training group. Thus only the mean raw score on the PPVT is shown for each treatment group in Table 13. The difference between the sentence training and control groups was not significant ($F < 1$, $df = 1/66$), but the vocabulary group scored significantly higher than the other two groups combined ($F = 32.78$, $df = 1/66$, $p < .001$). It can be seen that sentence training had virtually no effect on the preschoolers' ability to recognize items in the PPVT, but vocabulary training led to an advantage of about 14 words, on the average.

Effects of Training on Vocabulary Production. The PI transcripts were examined for the number of different target words produced by each subject. Various forms of a given word, such as singular and plural, were treated as tokens of a single word type. Not surprisingly, nearly all of the target word types counted in this way came originally from the PI rather than the PPVT. The preschool group means for this measure are shown in the second line of Table 13. There was a significant difference between the vocabulary group and the other two groups combined ($F = 8.17$, $df = 1/65$, $p < .01$), but not between the sentence and control groups ($F < 1$, $df = 1/65$). Thus the results for production of target words parallel the results for recognition.

Both vocabulary training and sentence training might conceivably have a transfer effect on the production of word types other than those

Table 13

Preschool Results:

Group Means on Five Dependent Measures from Posttest

(23 Subjects per Group)

Dependent measure	MSE	Treatment Group Means		
		Vocabulary	Sentence	Control
PPVT raw score	93.08	55.3	41.3	41.1
PI target word types	7.46	10.3	8.6[a]	8.0
PI other word types	1419.22	185.7	158.9[a]	170.2
SIT clearly correct	80.11	16.0	14.4	15.7
SIT clearly wrong	63.80	11.5	12.0	13.6

[a] _N_ = 22

which were selected for training. A count of the other word types produced by each preschool subject in the PI showed no significant difference between the sentence and control groups (F = 1.02, df = 1/65, p > .025) or between the vocabulary group and the other two groups combined (F = 4.83, .025 < p < .05). Thus only the target words were affected by vocabulary training, and there was no significant evidence that vocabulary or sentence training transferred to the production of other words.

Effects of Training on Sentence Imitation. Two global measures of proficiency in imitation were derived from the SIT. The first, called "clearly correct," was the number of sentences which the subject repeated essentially verbatim, with the allowance of a few very minor deviations. The second score, "clearly wrong," was the number of responses in which the subject obviously failed to preserve the meaning of the model sentence, either by omitting, adding, or substituting content words, or by altering basic grammatical relations, such as subject-object or modifier-head. In between the "clearly correct" and "clearly wrong" categories were a number of intermediate deviations, both standard and nonstandard, including changes in tense and number markers, grammatical transformations of the model sentence, and substitution of synonyms. Because many of these responses were ambiguous as to their correctness, they were excluded from the present analysis.

Although the "clearly correct" and "clearly wrong" measures are, of course, not entirely independent, the difference between them is important in principle because a child could receive low scores on both measures. It was expected that some subjects would translate many model sentences into their own dialects, without getting them clearly wrong. A lot of this translation did occur, but the "clearly correct" and "clearly wrong" scores also proved to be highly correlated (the r within treatment groups ranged from -.82 to -.84 on the posttest). However, the data from both measures are presented in the last two lines of Table 13. Neither variable showed a significant difference between the preschool subjects trained in sentence construction and the other two groups, nor were the contrasts between the vocabulary and control groups significant (F < 1, df = 1/66 in all cases). It was thought that the effect of sentence training might depend upon the skill of the particular trainer, or upon the pretest performance of the subject on the SIT. Supplementary analyses gave no evidence that either of these factors made any difference. In sum, the SIT showed no direct effects of sentence training, and no transfer effects of vocabulary training.

Kindergarten Results

Table 14 contains the group means for kindergarten subjects on three dependent measures. Since it was not possible to complete an analysis

- - - - - - - - - - - - - -

Insert Table 14 about here

- - - - - - - - - - - - - -

of the kindergarteners' Picture Interviews within the present project, only the results from the PPVT and the SIT are reported here.

Table 14

Kindergarten Results:

Group Means on Three Dependent Measures from Posttest

(22 Subjects per Group)

Dependent measure	MSE	Treatment Group Means		
		Vocabulary	Sentence	Control
PPVT raw score	46.86	68.7	56.2	55.3
SIT clearly correct	93.09	24.0	22.3	24.4
SIT clearly wrong	38.07	8.3	8.6	7.7

Effects of Training on Vocabulary Recognition. The kindergarten subjects who received vocabulary training produced a mean raw score on the PPVT which was significantly higher than the mean for the other two groups combined ($F = 52.09$, $df = 1/63$, $p < .01$). The subjects trained in vocabulary recognized about 13 more words, on the average, than the other two groups. The sentence training and control groups were not significantly different from each other ($F < 1$, $df = 1/63$). Thus sentence training had no significant effect on vocabulary recognition.

Effects of Training on Sentence Imitation. The kindergarten data from the SIT (See Table 14) were analyzed in the same way as the preschool data, with essentially the same results. Neither sentence training nor vocabulary training had a significant effect on sentence imitation performance. This outcome was obtained with both the "clearly correct" and "clearly wrong" scores. In all four of the planned comparisons, $F < 1$, $df = 1/63$.

Discussion

Vocabulary training had a significant positive effect on recognition and production of the target words which were taught. It is interesting to note that, on the PPVT, the mean raw score for the preschool vocabulary group was 55.3. Since a score of 60 represented the top of the range from which target words were selected, retention of the target words appears to have been quite good. As for the production of target words, the superiority of the preschool vocabulary group amounted to a difference of only two words. This may seem like a small gain, but the PI was conducted in such a way that the subject had to produce the target words more or less spontaneously, rather than in response to direct questions. It remains to be seen whether a similar effect occurred with the kindergarten subjects. These observations on the extent to which target words were recognized and produced suggest that vocabulary training may have effects which are significant educationally as well as statistically.

Although there was no significant evidence that vocabulary training transferred beyond the specific target words, the present results at least offer some encouragement in this direction. That is, the learning of target words -- a prerequisite for transfer -- did occur, and the transfer effect on the production of other words was nearly significant. The general effectiveness of vocabulary training might be increased by selecting target words with an eye toward their relationships to each other and to the needs of the child. The words in the present study were selected simply on the basis of the criterion tests, along with some data and intuition about the appropriate level of difficulty.

The positive effects of vocabulary training contrasted markedly with the lack of evidence that sentence training had any effect at all on the measures used in the present study. These findings imply that time devoted to early language training for disadvantaged black children is better spent on vocabulary than on sentence construction. This implication seems most relevant to short-term programs in which a limited amount of time is available for structured language activities. The present results are also consistent with the hypothesis that the syntactic aspects of language development are relatively impervious to direct instruction, perhaps because maturation plays an important role (Cazden, 1968). In this connection, it might even be argued that the present results favor a "difference model" over a "deficit model" with regard to the syntactic development of disadvantaged black children. That is, sentence training might have failed to have any effect because the subjects were already elaborating their sentences as much as one might expect for children their age. Conversely, a positive effect of sentence training might have implied the existence of a deficit or gap between readiness and actual achievement in the subjects' control of syntax.

Before the above conclusions are accepted, a number of questions ought to be raised about the negative results obtained with the SIT. The lack of significant group differences probably cannot be attributed to unreliability, since the test-retest correlations for the two SIT measures ranged from .80 to .92 within treatment groups. But it is quite possible to question the validity of the SIT, both as a specific measure of training effects in the present experiment and as a more general test of syntactic capability. First of all, the SIT seems less like the training situation than the PPVT or the PI. That is, the vocabulary subjects practiced pointing to pictures in response to words (as in the PPVT) and they practiced saying words in response to pictures (as in the PI), but the sentence subjects did not practice imitating sentences in the absence of visible referents (as in the SIT). Secondly, even though there is evidence that a child's imitations are related to the speech he produces and comprehends (Slobin, 1968), the relationship may not be close enough to warrant the use of imitation as a technique for assessing a child's general skill in manipulating syntactic structures. A detailed structural analysis of responses to the PI is in progress in order to provide a further test of training effects, and to check the validity of the SIT as a measure of ability or propensity to produce elaborated sentences.

Finally, it is possible that sentence training might have been more effective with different methods and/or content. Perhaps the target constructions could have been made more salient by means of a communication task which drew even more attention to the elaborated part of each sentence. Of course it is possible that the subjects were simply too young and egocentric (in the cognitive sense) to understand the need for elaboration in the service of communication. Also, the target constructions may not have been well chosen. During the training sessions, some constructions appeared to be quite easy for most of the subjects, while others seemed very difficult, perhaps because they were altogether beyond the subjects' grammatical competence. The syntactic analysis of PI responses may indicate that there was a poor match between the training materials and the capabilities of the subjects.

Tape recordings of the training session for two subgroups of preschool children were analyzed for additional hypotheses about the differences in training effects.[4] It was found that, on the average, children trained in sentence construction had fewer opportunities to respond within a training session than children trained in vocabulary. This is not surprising because, other things being equal, it takes longer to present or respond with a sentence, as opposed to a single word. In addition, however, the sentence training sessions were interrupted more often by disorderly behavior on the part of children within the group. These findings are based on a very small number of children, but they suggest that the sentence training involved fewer learning "trials" per subject, and that the task of learning sentence constructions was less meaningful than the task of learning vocabulary.

---

[4] This analysis was done by Mrs. Anne Smith.

In any case, the training and testing of sentence construction skills would appear to be much less straightforward than the training and testing of vocabulary. As indicated above, it is possible to recommend vocabulary training with some assurance of success. The same cannot be said with regard to training in sentence construction -- at least not the kind of sentence training which was attempted in this experiment.

Toward a Study of Sentence Imitation as a Technique for Assessing Language Ability

Paul R. Ammon

Given the recent concern among educators over the language of young children, it is surprising how little progress has been made in the development of techniques for assessing language ability in early childhood. Such techniques are needed especially for diagnosis and evaluation in connection with programs of early language training or enrichment. The training study reported in the preceding chapter might have been improved by the availability of a well-developed test for measuring ability in the area of syntactic elaboration. Instead, it was necessary to make use of an untried Sentence Imitation Test. The present chapter describes some further analyses of the Sentence Imitation Test data, with an eye toward future use of imitation as an assessment technique.

The psychometric use of sentence imitation, or "memory for sentences," goes back at least as far as Binet. Yet the imitation task is still largely unexplored as a way of measuring particular language abilities, rather than general intelligence. Sentence imitation might be expected to lend itself especially well to the assessment of syntactic skills. Syntax is an area of language which has generally been neglected by test makers; tests of vocabulary, or of auditory discrimination have been the primary measures of individual differences in language ability. Conversely, recent basic research in developmental psycholinguistics has been concerned almost exclusively with syntax—which at least underscores the fact that language acquisition involves more than the ability to distinguish speech sounds or the learning of vocabulary. But with a few exceptions (e.g., Lee, 1969), this theoretical interest in syntactic development has not yet been reflected in the way language ability is assessed for more applied purposes.

In connection with basic research on syntactic development, some progress has been made in devising techniques for assessing the young child's control of syntax in his native language (for a relatively comprehensive discussion of methods, see Slobin, 1967). This work has emphasized the assessment of language "competence"—the child's implicit knowledge of linguistic rules. The assessment of competence (in this technical sense) is seen as something of a methodological problem because of the possibility that a child's competence will either be over- or under-estimated by his"performance"—his actual behavior in producing speech or in responding to it. Under certain conditions, for instance, limitations in a child's memory capacity or in his perceptual and motor skills may prevent the child from manifesting the competence which he has. In developing a theory of language acquisition, it is important to distinguish such cases from ones in which a child simply has not yet learned a particular rule in his language. The common belief that young children can understand much more of language than they can produce has led to an emphasis on carefully contrived comprehension tasks as tests of linguistic competence, although production tasks have been found quite useful too.

The status of sentence imitation in relation to the distinction between competence and performance is not clear. There is good evidence that sentences are not simply treated as word lists or sound sequences in the imitation task, even by very young children (Slobin & Welsh, 1968). The types of errors which occur in imitation suggest very strongly that the child's linguistic competence enters into his processing of a model sentence. On the other hand, it is obvious that the imitation task involves both comprehension and production of speech, and therefore is influenced by a variety of performance factors. Thus the imitation task is analytically less "pure" than the methods which attempt to separate competence from performance, or to isolate comprehension and production. But this very complexity may turn out to be an advantage in using sentence imitation for assessment purposes in an applied setting. From a practical standpoint, a child's ability to coordinate linguistic competence and performance factors may be at least as important as any of these components taken separately.

Aside from this theoretical analysis, the sentence imitation task also recommends itself as an assessment tool because of its convenience. A large number of model sentences can be presented for imitation in a short period of time. Virtually all young children, at least down to age four, can understand the instruction to imitate. Many children seem to enjoy performing the task. The tester can, of course, systematically manipulate the model sentences in accord with those aspects of language ability which are of particular interest to him. The scoring of imitation responses can become rather complicated--as will be shown presently--but it does not seem unreasonable to expect that some relatively simple scoring methods will provide valid measures of language ability. Certainly the sentence imitation task would be more convenient than the collection and analysis of free speech samples. Some authors have suggested that samples of free speech ought to be included in a language assessment battery (e.g., Rosenberg, 1968), but these authors must not have considered the limitations on an assessment instrument if it is to be used by practitioners in the field.

In the development of a test, it is necessary but not sufficient to consider the issues of theoretical rationale and practical feasibility. A further, crucial step is the demonstration of validity. In the present case, it is a matter of specifying what sort of language ability is to be measured with the imitation task, and then of selecting an appropriate criterion against which to validate sentence imitation scores. As indicated above, sentence imitation seems best suited to the assessment of language ability in the area of syntax. Syntactic "elaboration" is the aspect of syntax which has attracted particular attention from educators concerned with the language of disadvantaged children (Bernstein, 1964, 1970; see also the preceding chapter of this report). Consequently, the strategy of the present study was to construct a Sentence Imitation Test which might assess a child's control of elaborated sentence constructions, and then to validate the child's performance on this test against measures of syntactic elaboration in relatively free speech, as derived from a Picture Interview. Both of these instruments--the Sentence Imitation Test and the Picture Interview--were introduced briefly in the preceding chapter and will now be discussed in much greater detail, both individually and in relation to each other.

It will soon become apparent that carrying out the above strategy has proved to be far from straightforward. The present chapter is, therefore, a report of work in progress.[1] However, the groundwork has been laid for a complete study of sentence imitation as a technique for assessing language ability in young children.

The Sentence Imitation Test

A list of 50 sentences was constructed for use in the Sentence Imitation Test (SIT). The number 50 was chosen arbitrarily as the near maximum number which might be administered to four- and five-year-old children in a single session. The complete list of sentences appears in Table 15.

- - - - - - - - - - - - - - -

Insert Table 15 about here

- - - - - - - - - - - - - - -

The primary goal of the list is to assess a child's ability to deal with some elementary forms of sentence complexity. It was decided to focus on complex noun phrases—especially cases in which the noun head is modified by some additional predication. The ability to use these forms of complexity in comprehension and production seems related to Bernstein's notion of an elaborated linguistic code. That is, the introduction of complexity elaborates on, or makes explicit, what the speaker is talking about. As a result of this focus on syntactic elaboration, many of the present sentences are somewhat longer than those which other investigators have used to elicit imitations from young children. Greater sentence length is quite consistent with the goal of assessing the child's control of syntax, however, since this control seems to reveal itself most readily in imitation when there is some strain on immediate memory.

A second objective of the sentence list in Table 15 concerns the detection of deviations from Standard English which are based in Black English dialect. The emphasis here is on syntactic and morphological deviations, rather than phonological or lexical differences. It is not always possible to distinguish the boundaries between these aspects of language in examining dialect differences, because differences on the level of phonology may produce grammatical deviations that are more apparent than real. For example, the apparent absence of an -ed inflection on a verb may actually be due to simplification of a word-final consonant cluster. In any case, the present emphasis on syntax and morphology led to the inclusion of many copular and auxilliary forms of be, other inflected verbs, possessive nouns, and pronouns of various sorts.

Construction of the Sentences

Before looking into the particular sentence constructions which were used, two characteristics of all the sentences ought to be noted. First, every sentence is a declarative statement. Although interrogative and imperative items may have an important place in language assessment, they were not included in the present list because it was necessary to limit the length of the list, and because some young children will attempt

---

[1] Research along the lines described here is being continued with the support of U.S. Office of Education Contract No. OEC-9-71-0039(508), "The Speech of Young Black Children: Individual and Group Differences in Syntax and Their Relationship to Reading."

Table 15

List of Items from the Sentence Imitation Test

1. The boy is eating his lunch.
2. The lady showed the people in the store some books.
3. The girl's father cut himself.
4. The teacher knows that the boys are sick.
5. The milk was in the cup that broke.
6. The teacher is wearing a green dress.
7. The doctor was called by an old man.
8. The boy's kitten is eating.
9. The picture in the book was funny.
10. The dog that caught the cat is drinking some water.
11. The children played a game until the teacher called them.
12. The clown the people watched was feeding a rabbit.
13. The toy is a doll that walks.
14. The girl who fell was crying.
15. The new teachers are pretty.
16. A man saw the car by the house.
17. The doctor is holding the kitten while the mother feeds it.
18. A book is under the teacher's chair.
19. The boy is the girl's brother.
20. The doctor brought the nice baby a toy.
21. The cat that ate was washing itself.
22. The children like their teacher.
23. The horses eat the apples that the boy brings.
24. The dinner was cooked by the lady in the kitchen.
25. The boys made some boats after they heard a story.
26. The girls found a book the teacher wanted.
27. The little birds are feeding themselves.
28. The mother made the girl who laughed some cookies.
29. The children told the teacher that some birds were singing.
30. The dolls are on a chair by the door.
31. The man who the children saw was drawing a picture.
32. The girl found her toys.
33. The clown's hat is red.
34. The happy children were playing.
35. The lady was a teacher at the school.
36. The rabbits were caught by the man's friend.
37. The lady likes the girl who is watching the baby.
38. The people are taking the apples that fell.
39. A clown is a funny man.
40. The girl who opened the door sees an airplane.
41. The shoes were in a brown box.
42. The lady in the house sings to herself.
43. The children who ran were hot.
44. The bird is building its nest.
45. The girl hit the boy when he took the candy.
46. The horses under the tree are sleeping.
47. The teacher called the man because she broke the chair.
48. A cow ate the lady's flowers.
49. A man gave the girl's dog some bread.
50. The windows were broken by the boy who ran.

to answer questions and obey commands, rather than imitate them. The second general remark pertains to vocabulary. To insure that all subjects in the target populations would be familiar with the words used in the sentences, a great majority of the words were drawn from the top 500 in Rinsland's (1945) frequency count for first graders. The remaining words came from the second 500 in Rinsland's list.

As for the construction of the individual sentences, consider first the set of sentences whose item numbers appear as entries in Table 16.

Insert Table 16 about here

This table describes the 32 combinations of eight basic sentence-types with four kinds of elaboration which could be added to each type. The underlined symbol in the formula for each basic sentence-type indicates which constituent contains the elaboration. Each of these 32 sentences contains the minimal number of content words needed to meet the specifications of Table 16. Thus, for example, all the relative clauses contain only an intransitive verb, in addition to the relative pronoun. A second subset of sentences is described by Table 17. These sentences

Insert Table 17 about here

are derived from two S + Vt + DO clauses in which one of the nouns occurs in both clauses. In the derived sentences, one clause is subordinated to the other, either as a relative clause, or by means of a subordinating conjunction plus a personal pronoun. In most cases, the repeated noun is replaced by a pronoun--either relative or personal--which refers to a head noun in the main clause. In two cases, however, the relative pronoun has been deleted in accordance with an optional rule of English grammar. Other than the type of subordination, the sentences in Table 17 vary according to which noun phrase the subordinate clause refers to (S or DO of the main clause), and according to whether the subject of the object noun has been deleted from the subordinate clause. Finally, a distinction is made between "confusable" and "non-confusable" constructions. A sentence is confusable if both the first and the second nouns are possible subjects of the final verb phrase; the intended predication is then marked only by the pronoun in the sentence, and not by selectional restrictions which apply to the nouns and the verb phrase in question.

Two other sentences involve subordination of the sort which is sometimes called indirect discourse. Item 4 has the construction S + Vt + that + S + Vc + Adj. Item 29 has the construction S + Vt + IO + that + S + Vi. Finally, four sentences contain third person possessive pronouns as modifiers of the object in the S + Vt + DO type of construction. These are items 1, 22, 32 and 44.

The order of sentences shown in Table 15 reflects an attempt to distribute the instances of each structural feature evenly throughout

Table 16

Item Numbers of 32 Sentences Representing Each Combination of Basic Sentence-Type X Kind of Elaboration

(See Table 15)

Basic Sentence-Type	Kind of Elaboration			
	Poss	Adj	Prep P	Rel
S + Vc + NP	19	39	35	13
S + Vc + Adj	33	15	9	43
S + Vc + Prep P	18	41	30	5
S + Vi	8	34	46	14
S + Vt + DO	48	6	16	38
S + Vt + IO + DO	49	20	2	28
Passive of S + Vt + DO	36	7	24	50
Reflexive: S + Vt + R	3	27	42	21

Key

Adj = adjective
DO = direct object
IO = indirect object
NP = noun phrase
Poss = possessive noun
Prep P = prepositional phrase: always indicates location
Rel = relative clause: always "who" or "that" + Vi
R = reflexive pronoun
S = subject
Vc = copular verb: is, are, was, or were
Vi = intransitive verb: } present and past, simple and progressive
Vt = transitive verb: }

Table 17

Item Numbers of 12 Sentences Derived from Two SVO Clauses

(See Table 15)

Type of Subordination	Features of Subordinate Clause				
	Refers to Main S			Refers to Main O	
	S Deleted		O Deleted	S Deleted	O Deleted
	Confus	NonCon			
Relative	10	40	31, 12[a]	37	23, 26[a]
Conjunctive	47	25	11	45	17

[a] Relative pronoun deleted.

the list, and to avoid interference between items. Toward these ends, the following procedure was used. The items were divided into four groups such that each group contained: (a) two items from each column and one from each row in Table 16; (b) three items from Table 17; and (c) one item with a possessive pronoun. Items 4 and 29 were added to the first and third groups respectively. Then, within a group, the items were randomly ordered, with the restriction that no two adjacent sentences would have any content words in common. This restriction was intended to reduce inter-item interference. Previous experience with lists of sentences suggested that much of the interference which occurs between items stems from lexical similarities between adjacent sentences.

Administration of the SIT

In the present study, the sentences listed in Table 15 were presented orally to individual subjects by the experimenters who were described in the preceding chapter. At the outset, the experimenter gave each subject the following instructions: "This is a remembering game. I am going to say something, and when I finish, you try to say just what I said. Let's try one." These instructions, and the first two items, were repeated as often as needed to insure that the child understood the task. (Little or no repetition was required in most cases.) Sentences 3 through 50 were repeated only once and only if the subject requested a repetition, gave no response, or appeared not to have heard the sentence. If the subject gave a minimal response (one or two words), the experimenter encouraged him to improve upon it (e.g., "can you remember the rest of it?"), but did not repeat the model sentence. The repetitions served primarily to check on the reliability of the initial response and were not considered further in scoring (with one exception which is noted in Appendix F).

Each model sentence was printed on a separate card and was pronounced clearly by the experimenter, with a normal speaking rate and intonation pattern. Although a tape recorded presentation would have permitted more uniformity of administration from one subject to the next, a "live" presentation seemed preferable in order to achieve maximum flexibility in the pacing of items, in the use of repetitions, and in the maintenance of the subject's attention. It seemed, too, that the visibly diminishing stack of sentence cards sometimes helped to sustain the subject's engagement with the task.

Transcription and Scoring of the SIT

The entire SIT session with each subject was tape recorded and later transcribed. The procedure for transcribing is not a trivial matter in this type of research. A certain amount of information is bound to be lost between tape and transcript. The investigator must therefore decide what sorts of information are crucial in the type of analysis which he anticipates doing. In the present study, the main desideratum was that syntactically and lexically relevant information be preserved in the transcript. Thus all morphemes that were realized in sound in some recognizable form were transcribed in the spelling of Standard American English. Variations in pronunciation were generally ignored. In addition, pauses and other hesitation phenomena were noted to some extent. Even so, it proved necessary to listen again to each tape at the time of scoring, due to the syntactic relevance of stress and intonation patterns. This relistening served also as a check on the accuracy of the original transcription.

Appendix E contains some general instructions for transcribing. These instructions apply to the Picture Interview as well as the SIT. Some additional guidelines apply only to the SIT, as follows. The experimenter's speech generally was not transcribed, except to take note of the deviations from, or repetitions of, the model sentence. The subject's responses were typed under the appropriate model sentence on a preprinted protocol.

The SIT transcripts were scored in terms of a coding scheme which treated the response to each model sentence as the unit of analysis. It is relatively easy to distinguish essentially verbatim imitations from all others, but it was felt that this kind of scoring would miss all of the rich information which might be found in the subjects' errors. On the other hand, the great variety of deviations from the model sentences necessitates some categorization of errors. As a first step in analysis of these errors, a six-category system was devised. Any given response was assigned to one, and only one, of the following categories.

1. Verbatim response
2. Small optional changes
3. Minor syntactic and/or lexical changes maintaining Standard English
4. Minor syntactic and/or lexical changes resulting in nonstandard English
5. Agnate sentences (transformations) and other responses with base structures similar to the model
6. Failure to maintain approximate base structure of the model

These categories are cumulative, in the sense that a response assigned to a higher-numbered category might also contain deviations associated with one or more lower-numbered categories. Categories 2 through 6 are illustrated by several examples in the complete scoring system presented in Appendix F. It should also be noted that the "clearly correct" and "clearly wrong" scores reported in the preceding chapter were based on categories 1 and 6 respectively.

The above set of categories is, admittedly, very crude, but it seemed advisable to postpone further refinement pending subsequent developments on two fronts. First, it would be useful simply to know the relative frequencies of responses in the present categories. If a particular category occurs very infrequently, then further refinement of the category would hardly be worthwhile. Second, revisions of the present scoring scheme might well be guided by the specific measures of elaboration which have yet to emerge from the Picture Interview data, and which will provide the criteria for validation of the SIT.

Preliminary Results from the SIT[2]

The SIT was administered as a pre- and posttest in the training study reported in the preceding chapter. Thus the data to be discussed in this section came from the same young black children who participated in the training experiment--69 preschoolers and 66 kindergarteners.

---

[2] The author gratefully acknowledges the help of Mary Sue Ammon, Pat Stohr, and Drs. Joel Levin and Nancy Suzuki in transcribing, scoring, and analyzing the SIT data.

The complexity of the SIT scoring system (see Appendix F) demanded a check on the extent of inter-scorer reliability in using the set of six categories. After considerable practice with the scoring system, a random sample of twenty protocols was drawn from the kindergarten pretest data and was scored independently by scorers A and B. Out of 1000 items, 29 were assigned to different categories by the two scorers (including clerical errors). Thus the scoring reliability for this sample and these scorers was 97.1% agreement. A sample of equal size from the preschool pretest yielded 98% agreement between scorers A and C. It seems fair to conclude that trained scorers can easily obtain 90% agreement or better using the present system of categories.

In order to determine the relative frequency of responses in the six categories, each subject was given a score equal to the number of responses which fell in each category. Table 18 contains the means and standard deviations of these scores for the two samples of subjects on the SIT pretest. On the average, the subjects did not repeat a large number of sentences verbatim, but neither did they fail very often to approximate the base structure and lexical content of the model sentence. Taken together, these two response categories (1 and 6) account for only slightly more than half of the total responses in both age groups. Of the remaining categories, number 4 (minor changes resulting in nonstandard English) was by far the most frequent, accounting for about 25 - 30% of the responses. This is not at all surprising, given the prevalence of Black English dialect among the children who were tested.[3]

- - - - - - - - - - - - - - -

Insert Table 18 about here

- - - - - - - - - - - - - - -

To reduce the number of variables in subsequent analyses, categories 3 and 5 were dropped from further consideration, due to their low frequencies. Category 2 responses seemed sufficiently close to verbatim that categories 1 and 2 were combined to form a new variable called "essentially correct." Category 4 was retained intact and was called "nonstandard approximation." (It should be noted, however, that some nonstandard approximations of the model occurred also in categories 3 and 5--see Appendix F.) Category 6 was retained, too, and was called "clearly wrong," as in the training study. The means and standard deviations for these three variables are shown for both the pre- and posttest in Table 19. In general, they account for about 85 - 90% of the total responses on the 50-item test.

- - - - - - - - - - - - - - -

Insert Table 19 about here

- - - - - - - - - - - - - - -

---

[3] It should not be assumed that all nonstandard deviations were based on Black English. A very rough estimate of the role of dialect in causing nonstandard deviations could be obtained by subtracting the data for white children from the present data on black children. Such data have been collected, but unfortunately they could not be analyzed within the present project.

Table 18

Means and Standard Deviations in Six Response Categories

for Pretest Administration of the Sentence Imitation Test

Response Category	Preschool (N=69)		Kindergarten (N=66)	
	Mean	S.D.	Mean	S.D.
1. Verbatim	8.9	7.3	16.9	9.5
2. Optional changes	2.9	1.8	4.7	2.9
3. Minor SE changes	3.9	2.3	4.5	2.0
4. Minor NSE changes	14.3	4.5	12.5	5.7
5. Agnate sentences	1.9	1.5	1.5	1.7
6. Failure to maintain base structure	17.8	9.1	9.9	6.6

Table 19

Means and Standard Deviations for Three Variables from the Sentence Imitation Test (Pre- and Posttest)

Variable	Preschool		Kindergarten	
	Mean	S.D.	Mean	S.D.
Essentially Correct: Pre	11.8	8.5	21.6	10.8
Essentially Correct: Post	15.4	8.8	23.6	9.5
Nonstandard Approx: Pre	14.3	4.5	12.5	5.7
Nonstandard Approx: Post	15.4	4.0	12.4	5.3
Clearly Wrong: Pre	17.8	9.1	9.9	6.6
Clearly Wrong: Post	12.4	7.9	8.2	6.1

The data described above have no general significance, because the occurrence of a particular kind of response depends greatly on the model sentence involved. On the pretest, for example, the number of subjects with a clearly wrong response to a single item ranged from 1 on item 34 to 89 on item 28 (total $\underline{N}$ = 135). Thus the relative frequency of different response types would be determined by the composition of the list of model sentences. The data do suggest, however, that the present list of sentences contains an appropriate range of difficulty for detecting individual differences among the subjects who were tested. It remains to be seen how stable an individual's score is on a particular variable, and how the variables relate to each other.

The intercorrelations of the three SIT variables from the pre- and posttests are shown in Table 20, for both the preschool and kindergarten samples. The first thing to note in this table is the high test-retest reliability for both the essentially correct and the clearly wrong scores ($\underline{r}$ = .85 to .91). These coefficients compare favorably with test-retest correlations of raw, MA, and IQ scores on the Peabody Picture Vocabulary Test, which range from .71 to .77. (The latter correlations are based only on those subjects who did not receive vocabulary training; otherwise the reliability coefficients would have been spuriously depressed by the effectiveness of intervening training for some subjects and not others. In these correlations, $\underline{N}$ = 46 and 44 for preschool and kindergarten respectively.) The test-retest reliability was somewhat lower for the nonstandard approximation scores, perhaps because of the heterogeneity of responses in this category, but perhaps also for some reasons to be mentioned below.

- - - - - - - - - - - - - - -

Insert Table 20 about here

- - - - - - - - - - - - - - -

A second interesting facet of Table 20 is the consistently high negative correlation between essentially correct and clearly wrong responses within the same testing ($\underline{r}$ = -.78 to -.82). While this result is perhaps not surprising, it was by no means a foregone conclusion. With the present samples of subjects, it was expected that some children--when their responses were not clearly wrong--would imitate the model sentence in an essentially correct form, while others would produce nonstandard approximations. Subjects in the latter group would tend to reduce the negative correlations between essentially correct and clearly wrong for the sample as a whole. The means and standard deviations in Table 19 indicate that some subjects in both samples produced a large number of nonstandard approximations, but apparently this did not have a marked effect on the correlation between essentially correct and clearly wrong scores.

The correlations between nonstandard approximations and the other two measures vary considerably according to the sample involved and the time of testing. The generally substantial negative correlations between nonstandard approximations and essentially correct responses (at the same time of testing) suggest that these two variables do, in fact, represent alternative strategies for dealing with the imitation task. This interpretation is bolstered further by the finding of little

Table 20

Intercorrelations of SIT Variables for Preschool (Below Diagonal) and Kindergarten (Above Diagonal)

Variable	1	2	3	4	5	6
1. Essentially Correct: Pre		.91	-.77	-.58	-.81	-.69
2. Essentially Correct: Post	.87		-.62	-.61	-.79	-.78
3. Nonstandard Approx: Pre	-.12	.04		.74	.33	.22
4. Nonstandard Approx: Post	-.43	-.47	.50		.20	.09
5. Clearly Wrong: Pre	-.79	-.79	-.41	.10		.85
6. Clearly Wrong: Post	-.73	-.82	-.35	-.06	.85	

or no relationship between a subject's propensity to produce nonstandard approximations and his ability to avoid being clearly wrong. Baratz (1969) obtained similar results with much older children.

The pretest data from the preschool children provide an exception to the above interpretation. There, the correlation between nonstandard approximation and essentially correct is negligible (-.12), while there is a moderate negative correlation of -.41 between nonstandard approximation and clearly wrong. But these data came from the youngest and least experienced subjects, who would have been doing well to produce even a nonstandard approximation of the model sentence. Essentially correct responses occurred less frequently than the other two types in the preschool pretest. Also, it will be recalled that the test-retest correlation on nonstandard approximations was relatively low for the preschoolers ($\underline{r}$ = .50).

In connection with the above correlations, it is interesting to note--in Table 19--that the number of nonstandard approximations is relatively constant, both between pre- and posttests within a sample, and even between the two samples. It may be that, while the number of nonstandard approximations does not seem to vary much with the frequency of clearly wrong responses, the set of particular items produced in nonstandard fashion does vary systematically. In other words, there might be a sort of "conveyor belt" effect, such that a particular item first elicits a clearly wrong response at the lowest level of ability, then a nonstandard approximation at some intermediate level, and finally an essentially correct response. For a child in the process of developing his ability, the number of items at the intermediate level on the conveyor belt would be fairly stable, while the numbers at the extremes would change. A second and equally plausible hypothesis would be that there is a fixed set of items which elicit nonstandard approximations to the extent that a given subject is so inclined, regardless of his ability to avoid clearly wrong responses. Obviously an item analysis would be very helpful here, but it was not possible to carry out such an analysis in the present project.

It would also be interesting to look at individual items from the SIT in relation to the variables specified in Tables 16 and 17--that is, type of elaboration, confusability, and the like. A preliminary inspection of the SIT protocols suggests that the children's imitations were influenced by these variables in ways that one might expect on the basis of previous research findings in psycholinguistics (See Bever, 1970, for a review).[4] But the present results are hardly definitive in this regard, because the sampling of sentences was not sufficient to control the effects of sentence variables which were not manipulated systematically in the present study. Consequently, this sort of analysis will not be discussed in any detail here. Suffice it to say that the SIT appears to have made contact with the same processes which have been revealed in other studies involving sentences, thereby giving the instrument some initial construct validity to go along with its face validity.

---

[4] The author thanks Carol Bell for her exploratory work in this area.

The high test-retest reliability of the "essentially correct" and "clearly wrong" scores stands out as the most striking thing about the SIT data. While these measures are admittedly global, as well as redundant with each other, their stability over a period of two to three months is impressive and offers some encouragement toward the further development of SIT measures. But, of course, it remains to be seen just what it is that the present SIT variables are measuring so reliably. Thus the next section presents a description and analysis of the criterion task, the Picture Interview.

The Picture Interview

This part of the study had as its goal the collection of speech samples to be compared with results from the Sentence Imitation Test. Toward this end, each child participated in a Picture Interview (PI) patterned after the interview used by Loban (1963). The PI consisted of a series of standard questions and pictures (see below) to which the subjects were encouraged to respond as freely as possible. This standard format for eliciting speech seemed preferable to more naturalistic observations because the aim was to attribute variance in performance to individual differences, rather than to differences between situations. On the other hand, the use of interview data undoubtedly involves certain risks, which ought to be discussed at the outset.

One common criticism (e.g., Labov, 1970) is that lower-class black children "clam up" when confronted by an adult in a one-to-one interview--especially if the adult is white. But in the present study, the great majority of children spoke rather freely and, to all appearances, were not intimidated by the interview situation. It may be that young children (four- and five-year-olds) are simply less sensitive to the pressures which cause some older children to be laconic in an interview. But aside from the child's inclination to talk or not, the speech samples collected by the present PI may be peculiar in other ways too. For one thing, much of a child's speech will, inevitably, consist of answers to questions--which may or may not be representative of the child's speech in other contexts. Also, when a child is asked to describe a picture which the interviewer obviously can see for himself, it is hard to say what consequences this might have for the way in which speech is used.

The general implication of factors such as those outlined above is that the investigator must be very careful not to confuse a child's language ability with his reaction to being interviewed. A similar principle applies to the analysis of speech in any other setting as well. If nothing else, the present PI data provide a basis for developing analytic techniques which could be used on speech samples from any number of situations--and a great variety of situations will have to be sampled for a complete study of language ability and its assessment in young children.

Content and Administration of the PI

The PI consisted of two parts: a "warm-up" portion and a picture portion. In the warm-up, the child was questioned briefly on a series of standard topics, as follows.

> Can you tell me your whole name?
> How old are you?
> Do you have any brothers or sisters?
> What have you been doing in school today?
> What's your favorite thing to do?
> What do you like to watch on TV?
> Can you tell me about one time when you were sick?
> Do you have a dog or any kind of pet at your house?

Additional probes and follow-up questions on these topics depended, of course, on the child's responses. The general objective was to engage the child in relaxed conversation and to encourage speech on his part. As the name "warm-up" implies, this portion of the interview was intended primarily to establish rapport. But some children produced proportionately large amounts of speech during the warm-up, so it is regarded as being no less important than the picture part, which follows.

In the second portion of the PI, the child was shown a series of six pictures, one at a time. Each picture was introduced with an open-ended instruction such as "Tell me everything you can about this picture." Appropriate probes and follow-up questions were, again, based on the child's initial responses. The six pictures were collected from assorted magazines, calendars, and other sources, and had been pretested for their conversational value with young children.

In the present study, the PI was always administered individually, following the SIT, but on a different day. The interview schedule was planned with a twenty-minute session in mind, but this varied considerably from one subject to the next. Each PI session was tape recorded in its entirety.

Transcription and Scoring of the PI

The importance of a systematic approach to transcription has already been discussed in connection with the SIT. The reader is referred again to Appendix E, the guidelines for transcribing tapes in the present study. Perhaps two additional points should be emphasized with particular reference to the PI. First, an accurate record of the interviewer's speech, as well as the child's, was deemed as essential for later interpretation of the child's speech. Second, in an effort to guarantee accuracy of transcription so far as possible, each tape was heard by at least two transcribers. More difficult passages (and there were many) were heard by three or four people, who conferred as to the most likely interpretation of what was being said. Needless to say, this is a very time consuming procedure, but also a necessary one if the research is to accomplish its ultimate objectives.

When the present study began, it seemed as if the scoring of PI transcripts would be a relatively straightforward matter. Other researchers had done structural analyses of interview speech samples and had offered their findings as support for Bernstein's (1964) theory about social class differences in the use of an elaborated code (e.g., Loban, 1963; Hess & Shipman, 1965; Williams & Naremore, 1969b). It seemed, then, as if measures of syntactic elaboration were readily available and, when applied to the present PI transcripts, would provide appropriate criterion variables for validation of the SIT. Upon closer scrutiny, however, it became apparent that these older measures were fraught with problems which would necessitate the development of new methods of analysis. A brief review of the shortcomings of the more traditional measures is appropriate here, because it will elucidate both the importance and the magnitude of the task which was undertaken in the present study when it was decided to devise a new scoring system. These remarks will be focussed on the problem of measuring elaboration in the speech of lower-class black children in particular.[5]

At the outset, it is important to heed Bernstein's (1970) warning that there is no inherent connection between a speaker's dialect and his use of the elaborated code. Most researchers have recognized that one cannot simply count dialect deviations from Standard English as evidence of a deficiency in elaboration. There is no reason to believe that the linguistic system which generates such deviations is inferior and is therefore a deficit--except as a social liability when black people are dealing with the white establishment. Unfortunately, however, it sometimes is difficult to untangle dialect phenomena from elaboration per se. This problem has not been recognized sufficiently by researchers in the past.

Consider a very simple, but common measure in language behavior research--mean sentence length. There is a certain logic in using sentence length to measure elaboration. The number of words in a sentence does increase with the addition of elaborative words, phrases, or clauses--other things being equal. The problem is that other things are not equal when speakers of different dialects are being compared. To take an obvious example, copular forms of be are often realized as a zero morpheme in Black English, as in she my best fiend. Such "omissions" do not reduce syntactic complexity, but they do reduce the mean number of words per sentence in the speech of a lower-class black child.

In at least one study (Loban, 1963), the deletion of copular verbs in Black English seems to have affected another syntactic measure associated with Bernstein's "restricted" code--the frequency of "incomplete" or "partial" sentences. The sentence she my best friend contains no verb (at least on the surface) and it might therefore be scored as incomplete. By this criterion, a lower-class child speaking Black English would appear to have more incomplete sentences than his middle-class counterpart--

---

[5] The remarks which follow were included in a paper presented at the annual meeting of the California Educational Research Association, San Diego, April 30, 1971.

other things being equal. In reality, however, a sentence with the copula deleted is no less complete than a sentence with the copula contracted, as in *she's my best friend*.

There is a need, then, for measures of elaboration that are more direct and less crude than sentence length or completeness.[6] A count of elaborative elements themselves seems quite direct. The traditional favorite here is the number of subordinate or dependent clauses. A higher frequency of subordinate clauses would indicate a more elaborative type of speech. But again the researcher may run into snags involving dialect differences. In one of the PI sessions, a preschool child said *there's a girl live in a house with a cherry tree*. This sentence appears to contain a relative clause, even though it lacks the relative pronoun required in Standard English. Thus one must at least be sensitive to the ways in which subordinate clauses are formed in non-standard dialects.

But the counting of subordinate clauses raises an even more fundamental question. What is so special about the *clause* as an elaborative element? What about infinitive and participial phrases, for instance? Even when researchers have tried to incorporate these other constructions in their measures of subordination, they still have given more weight to clauses (Loban, 1967). While this is not, strictly speaking, a dialect problem, there may be stylistic differences closely related to dialect, such that one group of speakers prefers to elaborate more often with clauses. But is it really less restricted to say *the boy who is climbing the tree*, as opposed to *the boy climbing the tree*? The second construction not only lacks a dependent clause but also has fewer words in it. Instead of counting clauses or words, why not simply count different types of elaboration and report any differences in frequency between groups of supjects? This can be done without assigning arbitrary weights to various constructions, on the assumption that some are better than others.

It is important to avoid a confusion between the syntactic and lexical aspects of elaboration. Some researchers have counted the number of *uncommon* adjectives and adverbs as an index of elaboration (e.g., Hess & Shipman, 1965), but this seems to reflect the richness of the speaker's vocabulary, rather than the use of modification per se. On the other hand, it is also important to examine the lexical content of sentences which are being scored for syntactic elaboration. Sometimes the lexical content has structural implications. The verb *put*, for example, requires not only a direct object but also a locative word or phrase. That is, one always talks about putting something some*where*. In other places, a locative phrase might be considered an instance of elaboration, but with *put* it is just a necessary part of the sentence.

It is conceivable that certain groups of subjects, speaking on certain topics, will differ with regard to the frequency of certain lexical contents. There is a very blatant example which seems nonetheless to have been overlooked. It is the response I don't know, which occurs with some regularity in children's interviews, perhaps more often with lower-class children. If I don't know is treated as just another sentence, one finds that it contains very few words, no dependent clauses, and no modifiers—uncommon or otherwise. But the occurrence of I don't know probably tells us more about a child's readiness to answer questions than about his ability or inclination to produce elaborated speech.

---

[6] In the study by Williams and Naremore (1969b), "a quantitative description of syntactic elaboration was obtained by use of a modified immediate constituents procedure which provides coding of the structural divisions of English sentences." This procedure has the advantage of identifying the part of a sentence in which elaboration occurs, but it also amounts to counting the number of words per sentence constituent, and therefore it has the same shortcomings as a count of the words in a whole sentence.

An analysis of syntactic elaboration must consider not only the content of a sentence, but also its context. Only one aspect of context will be mentioned here, but it is potentially a very important one. In an interview, the immediate context for many of a child's sentences is a direct question or some other request for information. Other sentences expand upon these immediate responses to questions, or they introduce information which has not been requested by the interviewer. There is some evidence that middle-class children tend to produce more of these expansions and spontaneous remarks (Williams & Naremore, 1969a). In other words, immediate answers to questions probably account for a greater proportion of the sentences produced by lower-class children. This may have more to do with the social psychology of interviews than with a child's language ability.

The point is that direct responses to questions and more spontaneous remarks ought to be analyzed separately in a study of syntactic elaboration, because there may be systematic structural differences between sentences produced in these two contexts. For one thing, the immediate answer to a question tends to be short--often elliptical--because some information has already been made explicit in the question. Furthermore, to the extent that the questions are about the person being interviewed, his immediate answers are likely to be permeated by the personal pronouns I and me, which do not lend themselves to modification by adjectives or adjective phrases. On both of these counts, then, the sentences of a lower-class child would come out looking less elaborated, so long as the researcher ignored the contexts in which they occurred.

The main thrust of the preceding discussion is that measures of syntactic elaboration which have been used in the past are too vulnerable to the influence of other social class or ethnic differences in language behavior--differences which have little if anything to do with the ability to produce elaborated speech. This criticism has implications not only for the immediate task of investigating the validity of the SIT, but also for the status of the basic assumption that lower-class black children tend to be relatively deficient in their ability to produce elaborated speech. At the beginning of the present project, the SIT and the PI were administered to samples of white middle-class children, in addition to the black children who were described earlier. The original objective was to compare the two groups in order to shed further light on differences in complexity of syntax which were already assumed to exist. Now it seems that such a comparison might be used for the purpose of reconsidering the whole assumption regarding lower-class deficits in syntactic elaboration. This part of the research program has yet to be carried out.

In order to circumvent the problems reviewed above, work began on a new system for describing and scoring the variety and complexity of syntax in the PI. One component of the system treats the verb-complement pattern as the basic unit of analysis, while a second component is designed for the analysis of noun phrases. This approach was suggested by the sociolinguistic work of Shuy, Wolfram, and Riley (1967), which in turn was based primarily on a linguistic analysis of English by Gleason (1965). Both components of the present system are outlined in Appendix G. The entire system has undergone several revisions and is still subject to further changes when new problems are encountered in the raw data. However, some analyses have been done with regard to verb-complement units, so the remaining discussion will focus on that part of the system.

The system for coding verb-complement units has a few features which ought to be pointed out briefly before the preliminary analyses are presented. (For more detailed information about the system and its linguistic background, the reader is referred to the work of Gleason and of Shuy _et al_. cited above.) For one thing, the system distinguishes 18 different types of verb-complement pattern, so that the present structural categories will be more homogeneous than those used by previous investigators. This is important because the amount and kind of elaboration may depend to some extent on the type of basic pattern involved. It also permits a more fine grained analysis of the variety of syntactic constructions used by a given speaker. A second feature of the system is that it codes the occurrence and the type of deletions in verb-complement units. This should allow for the discrimination of deletions associated with dialect from other kinds of deletions. The system also codes the grammatical function of each verb in such a way that dependent clauses are just one of the many subordinate uses of verbs that can be identified. Another feature is the coding of the environment in which each verb-complement unit occurs, e.g., first clause in response to a question, request, or command from the interviewer. Environment is further specified with regard to the part of the interview in which a unit occurs. Finally, certain stereotyped units, such as _I don't know_, are clearly marked and may be segregated from the remaining speech for purposes of analysis.

One more issue deserves comment here. It might be argued that the ultimate criterion of language ability lies in some measure of communication effectiveness, not syntactic elaboration. Indeed, too much elaboration may be said to impede communication (Labov, 1970). Thus a structural analysis of speech must eventually be coupled with a functional analysis. But for the time being, it seems reasonable to assume that measures of syntactic complexity or elaboration provide good indices of language ability, at least with four- and five-year-old children.

Preliminary Results from the PI[7]

The time spent on transcribing the PI and developing a scoring system for it precluded any extensive analyses with the present project. However, the pre- and posttest transcripts for 12 preschool subjects have been completely coded in terms of the scheme for analyzing verb-complement units, so some preliminary results can be reported here. This sample includes four children--two boys and two girls--from each of the training groups described in the preceding chapter. The small size of this sample should be kept in mind throughout the following discussion. The present analyses merely serve in illustrative purpose, since they barely scratch the surface of the possibilities inherent in the PI scoring system.

---

[7] The author is grateful to Mary Sue Ammon, Judy Harker, Phyllis Koppelman, and Sylvia Zuck for making this section and the next possible.

Table 21 contains the means and standard deviations for five variables derived from the scoring of verb-complement units (see Appendix G). The first variable is a very rough measure of elaboration. It is the

- - - - - - - - - - - - - - - -

Insert Table 21 about here

- - - - - - - - - - - - - - - -

percentage of verb-complement units in which the basic pattern has been expanded by the use of one or more modifying elements. It is interesting to note that over half of the verb-complement units involved some kind of expansion. The second variable--the percentage of units in which one or more elements in the basic verb-complement pattern have been deleted--indicates a subject's tendency to abbreviate his utterances. It is undoubtedly too gross a measure in its present form, because it includes all kinds of deletions, ranging from those which are obligatory in Standard English, through those which are optional, to those which are prohibited but may be consistent with a nonstandard dialect. The first two variables are conceptually independent, since a verb-complement unit can have both expansions and deletions at the same time. The third variable is the percentage of verb-complement units in which the verb was not the main verb in a sentence. Thus it indicates the extent to which verb-complement units were used to expand on a basic pattern, although it also includes the complements of catenative verbs, which are not necessarily instances of elaboration. The fourth variable is the number of different verb-complement patterns (out of 18 types) which were used at least once in a given transcript. This measure reflects the extent to which a subject makes full use of the various syntactic constructions which are possible. Since this is clearly too crude a variable to serve its purpose, a more refined measure of syntactic variety is currently being developed. The fifth and final variable is simply the total number of verb-complement units produced by the subject in a given PI session. Aside from indicating the baseline from which percentage scores were derived, it is a measure of the subject's total output of speech in the interview.

The variables listed in Table 21--especially the first four--were selected for their relevance to the notion of syntactic elaboration. The relationships among them are therefore of interest. The intercorrelations of the five variables are displayed in Table 22, with the test-retest correlations shown in parentheses along the main diagonal. Except for percentage without deletions, the test-retest reliability of the measures

Table 21

Means and Standard Deviations for Five Variables from Picture Interview for Preschool Subjects (N = 12)

Variable	Pretest		Posttest	
	Mean	S.D.	Mean	S.D.
1. % Expansion	55.2	12.6	60.3	9.1
2. % No Deletion	49.3	20.3	54.2	16.6
3. % Not Main Verbs	17.0	5.3	21.0	6.3
4. No. Different Patterns	11.0	1.9	12.1	1.5
5. Total V-C Units	139.4	71.8	169.2	58.5

is only fair to poor. The relatively high reliability of deletions may be due in part to those deletions associated with dialect. Low reliability is a cause for concern here, because the goal is to measure a relatively stable personal trait, i.e., language ability. Several explanations may be proposed for the present reliability data, although none of them has yet been explored further.

- - - - - - - - - - - - - - -
Insert Table 22 about here
- - - - - - - - - - - - - - -

First, it seems unlikely that the generally mediocre level of reliability is a peculiarity of the 12 children involved; their test-retest correlation for SIT "essentially correct" scores was .88. It is possible, however, that some of the subjects were changed by their exposure to training, and that these changes have manifested themselves in the PI, even though they did not do so on the SIT. A test of training effects with the full sample would be most helpful here. Another, strong possibility is that these relatively crude PI measures need further refinement. Perhaps reliability will be improved when factors like environment and type of verb-complement pattern are taken into account. On the other hand, the present speech samples may simply be inadequate, in which case no amount of psychometric refinement will produce reliable measures. It would be disappointing to arrive at this last conclusion, but it would also be an important finding.

Notwithstanding the problem of reliability, the intercorrelations of PI measures within a given testing still hold some interest. In the pretest, all five variables correlate with each other at a fairly high level. Some of the highest correlations involve the total number of verb-complement units, suggesting that all of the correlations may reflect some sort of general fluency factor. With total units partialled out, the remaining correlations might be reduced markedly. An outcome of that sort would not necessarily invalidate all or even any of the present variables as measures of elaboration. It might mean that elaboration is multidimensional--that different subjects have different styles of elaboration. In any case, more refined measures may, again, hold the key, by taking sheer volume of output into account through scoring rather than doing it statistically.

With just one exception, the correlations for the posttest are lower than their counterparts for the pretest. This is especially true of those correlations which involve the two variables with the lowest test-retest reliability. Thus the lower intercorrelations on the posttest may reflect the same factors which were discussed in connection with reliability.

The preliminary results from the PI have raised more questions than they have answered, but it is fairly clear now what must be done to answer some of these questions. It is also clear that this type of analysis must be done with larger samples of subjects before any firm conclusions can be drawn.

Table 22

Intercorrelations of Five Variables from the Picture Interview, Pretest (Above Diagonal) and Posttest (Below Diagonal)

Variable	1	2	3	4	5
1. % Expansion	(.25)	.85	.65	.63	.80
2. % No Deletion	.46	(.79)	.81	.66	.69
3. % Not Main Verbs	.21	.68	(.23)	.70	.67
4. No. Different Patterns	.49	.54	.13	(.47)	.90
5. Total V-C Units	.62	.84	.50	.52	(.56)

The Sentence Imitation Test as a Predictor of Performance in the Picture Interview

This chapter ends with one small example of the kind of analysis toward which the whole study is aimed. Table 23 shows the correlations of SIT "essentially correct" scores with the five variables derived from the PI. For the sake of comparison, the table includes a parallel set of correlations with mental age scores from the Peabody Picture Vocabulary Test. Only the pretest correlations are presented, because the posttest correlations would be distorted by effects of training on Peabody scores, and possibly on the PI scores as well.

Insert Table 23 about here

In every row of Table 23, the SIT score is a better predictor of PI performance than is the Peabody, which has very little correlation with the PI. The uniformity of these results is, at least in part, a function of the high correlations among the PI measures themselves. Nonetheless, the data in Table 23 suggest (a) that the SIT does, in fact, measure a child's ability to produce syntactically elaborated speech, and (b) that this ability does not overlap a great deal with the kind of ability that is measured by the Peabody. The last point is supported further by a pretest correlation of .48 between the SIT and the Peabody for the small sample, and a correlation of .47 for the preschool sample as a whole.

These results, and those reported earlier, show promise. Much work has been done in laying the foundation for complete study of sentence imitation as an assessment technique. The effort seems more worthwhile than ever now, but much work remains to be done.

Table 23

Intercorrelations of Sentence Imitation and Peabody Vocabulary Scores with Five Variables from Picture Interview Pretest

Variable	SIT Essentially Correct	Peabody MA
1. % Expansion	.36	-.19
2. % No Deletion	.52	-.02
3. % Not Main Verbs	.34	.15
4. No. Different Patterns	.55	.16
5. Total V-C Units	.40	.08

References

Ammon, P. R., & Ammon, M. S. Effects of training young black children in vocabulary vs. sentence construction. Paper presented at the annual meeting of the American Educational Research Association, Minneapolis, March, 1969.

Ammon, P. R., & Ammon, M. S. Effects of training black preschool children in vocabulary vs. sentence construction. Journal of Educational Psychology, in press.

Baratz, J. C. A bi-dialectal task for determining language proficiency in economically disadvantaged Negro children. Child Development, 1969, 40, 889-902.

Bereiter, C., & Engelmann, S. Teaching disadvantaged children in the preschool. Englewood Cliffs, N.J.: Prentice-Hall, 1966.

Bernstein, B. Elaborated and restricted codes: Their social origins and some consequences. In J. J. Gumperz and D. Hymes (Eds.), The ethnography of communication. American Anthropologist, 1964, 66, Part 2, 55-69.

Bernstein, B. A sociolinguistic approach to socialization: With some reference to educability. In F. Williams (Ed.), Language and poverty, Chicago: Markham, 1970.

Bever, T. The cognitive basis of linguistic structure. In J. R. Hayes (Ed.) Cognition and the development of language. New York: Wiley, 1970.

Carry, L. R. Interaction of visualization and general reasoning abilities with instructional treatment in algebra. Unpublished doctoral dissertation, Stanford University, 1967.

Cazden, C. B. Environmental assistance to the child's acquisition of grammar. Unpublished doctoral dissertation, Harvard University, 1965.

Cazden, C. B. Subcultural differences in child language: An interdisciplinary review. Merrill-Palmer Quarterly, 1966, 12, 185-219.

Cazden, C. B. Some implications of research on language development for preschool education. In R. D. Hess and R. M. Bear (Eds.), Early education. Chicago: Aldine, 1968.

Cronbach, L. J. How can instruction be adapted to individual differences? In R. M. Gagne (Ed.), Learning and individual differences. Columbus: Merrill, 1967.

Cronbach, L. J. Intelligence? Creativity? A parsimonious reinterpretation of the Wallach-Kogan data. American Educational Research Journal, 1968, 5, 491-511.

Cronbach, L. J., & Snow, R. E. Individual differences in learning ability as a function of instructional variables. Final report, U.S.O.E. contract # OEC-4-6-061269-1217, March 1969.

Crowder, N. A. Automatic tutoring by intrinsic programming. In A. A. Lumsdaine & R. Glaser (Eds.), Teaching machines and programmed learning. Washington, D. C.: National Education Association, 1960.

Davidson, R. E., & Adams, J. F. Verbal and imagery processes in children's paired-associate learning. Journal of Experimental Child Psychology, 1970, in press.

Deutsch, M. The disadvantaged child and the learning process. In A. H. Passow (Ed.), Education in depressed areas. New York: Teachers College, Columbia University, 1963.

Dickie, J. P. Effectiveness of structured and unstructured (traditional) methods of language training. In M. A. Brottman (Ed.), Language remediation for the disadvantaged preschool child. Monographs of the Society for Research in Child Development, Serial No. 124, 1968, 33, No. 8.

Dilley, M. G. & Paivio, A. Pictures and words as stimulus and response items in paired-associate learning of young children. Journal of Experimental Child Psychology, 1968, 6, 231-240.

Dunn, L. M. Expanded manual: Peabody Picture Vocabulary Test. Minneapolis: American Guidance Service, Inc.,1965.

Elkind, D. Piagetian and psychometric conceptions of intelligence. Harvard Educational Review, 1969, 39, 319-337.

Fredrick, W. C., Blount, N. S., & Johnson, S. L. A comparison of verbal statement, symbolic notation and figural representation of grammar concepts. Technical Report No. 64, Wisconsin Research and Development Center for Cognitive Learning, Madison, October 1968.

Glasman, L. D. A social-class comparison of conceptual processes in children's free recall. Unpublished doctoral dissertation, University of California, Berkeley, 1968.

Gleason, H. A. Linguistics and English grammar. New York: Holt, Rinehart and Winston, 1965.

Hays, W. L. Statistics for psychologists. New York: Holt, Rinehart and Winston, 1963.

Hess, R. D., & Shipman, V. C. Early experience and the socialization of cognitive modes in children. Child Development, 1965, 36, 869-886.

Irwin, M. H., Gerdes, B. J., & Rohwer, W. D., Jr. SES, age, and facilitation of free recall learning. Paper presented at the annual meeting of the American Educational Research Association, New York, 1971.

Jensen, A. R. Varieties of individual differences in learning. In R. M. Gagne (Ed.), Learning and individual differences. Columbus: Merrill, 1967.

Jensen, A. R. How much can we boost IQ and scholastic achievement? Harvard Educational Review, 1969, 39, 1-123.

Jensen, A. R., & Frederikson, J. Social class differences in free recall of categorized and uncategorized lists. In A. R. Jensen and W. D. Rohwer, Jr. An experimental analysis of learning abilities in culturally disadvantaged children. Final Report, U.S.Office of Economic Opportunity, Contract # OEO 2404. July, 1970.

Jensen, A. R., & Rohwer, W. D., Jr. Syntactical mediation of serial and paired-associate learning as a function of age. Child Development, 1965, 36, 601-608.

King, F. J., Roberts, D., & Kropp, R. P. Relationship between ability measures and achievement under four methods of teaching elementary set concepts. Journal of Educational Psychology, 1969, 60, 244-247.

Klaus, R. A., & Gray, S. W. The Early Training Project for disadvantaged children: A report after five years. Monographs of the Society for Research in Child Development, 1968, 33, No. 4.

Kohlberg, L. Early education: A cognitive-developmental view. Child Development, 1968, 39, 1013-1062.

Labov, W. The logic of nonstandard English. In F. Williams (Ed.), Language and poverty. Chicago, Markham, 1970.

Lee, L. Northwestern syntax screening test. Evanston, Ill.: Northwestern University, 1969.

Levin, J. R., Rohwer, W. D., Jr., & Cleary, T. A. Individual differences in the learning of verbally and pictorially presented paired associates. American Educational Research Journal, 1971, 8, 11-26.

Loban, W. The language of elementary school children. (Research Report No. 1). Champaign, Illinois: National Council of Teachers of English, 1963.

Loban, W. Language ability: Grades ten, eleven, and twelve. Final Report, Project # 2387, U. S. Office of Education, Bureau of Research, 1967.

Messick, S. The criterion problem in the evaluation of instruction: Assessing possible, not just intended outcomes. Report No. 22, Center for the Study of Evaluation of Instructional Programs, UCLA, May 1969.

Milgram, N. A. Retention of mediation set in paired-associate learning of normal children and retardates. Journal of Experimental Child Psychology, 1967, 5, 341-349. (a)

Milgram, N. A. Verbal context versus visual compound in paired-associate learning by children. Journal of Experimental Child Psychology, 1967, 5, 597-603. (b)

Milgram, N. A., & Riedel, W. Verbal context and visual compound in paired-associate learning of mental retardates. American Journal of Mental Deficiency, 1969, 73, 755-761.

Osborn, J. Teaching a teaching language to disadvantaged children. In M. A. Brottman (Ed.), Language remediation for the disadvantaged preschool child. Monographs of the Society for Research in Child Development, Serial No. 124, 1968, 33, No. 8.

Paivio, A., & Yarmey, A. D. Pictures versus words as stimuli and responses in paired-associate learning. Psychonomic Science, 1966, 5, 235-236.

Raph, J. B. Language development in socially disadvantaged children. Review of Educational Research, 1965, 35, 389-400.

Rinsland, H. D. A basic vocabulary of elementary school children. New York: Macmillan, 1945.

Rohwer, W. D., Jr. Social class differences in the role of linguistic structures in paired-associate learning: Elaboration and learning proficiency. Final report, U.S. O. E. Basic Research Project # 5-0605, Contract # OE6-10-273. November 1967.

Rohwer, W. D., Jr. Mental elaboration and proficient learning. In J. P. Hill (Ed.), Minnesota symposia on child psychology, Vol. 4, Minneapolis: University of Minnesota Press, 1970.

Rohwer, W. D., Jr. Adolescence: Prime time for education. Harvard Educational Review, in press.

Rohwer, W. D., Jr., & Ammon, M. S. Locus of facilitation in noun-pair learning. Paper presented at the annual meeting of the Western Psychological Association, 1968, San Diego.

Rohwer, W. D., Jr., & Ammon, M. S. Elaboration training and learning efficiency in children. Journal of Educational Psychology, in press.

Rohwer, W. D., Jr., Ammon, M. S., Suzuki, N., & Levin, J. R. Population differences and learning proficiency. Journal of Educational Psychology, 1971, 62, 1-14.

Rohwer, W. D., Jr., & Levin, J. R. Elaboration preferences and differences in learning proficiency. In J. Hellmuth (Ed.), Cognitive studies: deficits in cognition, Vol. 2. New York: Brunner/Mazel, Inc., 1971.

Rohwer, W. D., Jr., Lynch, S., Levin, J. R., & Suzuki, N. Pictorial and verbal factors in the efficient learning of paired associates. Journal of Educational Psychology, 1967, 58, 278-284.

Rohwer, W. D., Jr., Lynch, S., Suzuki, N., & Levin, J. R. Verbal and pictorial facilitation of paired-associate learning. *Journal of Experimental Child Psychology*, 1967, 5, 294-302.

Rosenberg, S. Problems of language development in the retarded: A discussion of Olson's review. Paper prepared for the Peabody-NIMH Conference on Social-Cultural Aspects of Mental Retardation, Nashville, Tennessee, June, 1968.

Semler, I. J., & Iscoe, I. Comparative and developmental study of the learning abilities of Negro and white children under four conditions. *Journal of Educational Psychology*, 1963, 54, 38-54.

Shuy, R. W., Wolfram, W. A., & Riley, W. K. *Linguistic correlates of social stratification in Detroit speech*. Final Report, Cooperative Research Project 6-1347, U.S.O.E., 1967.

Slobin, D. I. (Ed.), *A field manual for cross-cultural study of the acquisition of communicative competence.* Berkeley: University of California, 1967.

Slobin, D. I. Imitation and grammatical development in children. In N. S. Endler, L. R. Boulton, and H. Osser (Eds.), *Contemporary issues in developmental psychology*. New York: Holt, Rinehart and Winston, 1968.

Slobin, D. I., & Welsh, C. A. Elicited imitations as a research tool in developmental psycholinguistics. Working Paper No. 10, Language-Behavior Research Laboratory, University of California, Berkeley, 1968.

Stevenson, H. W., & Siegel, A. Effects of instructions and age on retention of filmed content. *Journal of Educational Psychology*, 1969, 60, 71-74.

Suppes, P., & Morningstar, M. Computer-assisted instruction. *Science*, 1969, 166, 343-350.

Wallach, M. A., & Kogan, N. *Modes of thinking in young children*. New York: Holt and Co., 1965.

Williams, F., & Naremore, R. C. On the functional analysis of social class differences in modes of speech. *Speech Monographs*, 1969, 36, 77-102. (a)

Williams, F., & Naremore, R. C. Social class differences in children's syntactic performance: A quantitative analysis of field study data. *Journal of Speech and Hearing Research*, 1969, 12, 778-793. (b)

Yuille, J. C., & Paivio, A Imagery and verbal mediation instructions in paired-associate learning. *Journal of Experimental Psychology*, 1968, 78, 436-441.

Appendix A

Objectives and Procedures of Daily Treatment Sessions

Day 1 - Tutorial

Purpose of Activity		Materials
Have child get to know trainer+situation. Promote verbal interaction. Get children to perform difficult task in non-pressure situation.	1. Have children try puppets on and tell them they (the puppets) are going to get a chance to play the television game like the children did before.	2 sets of 2 puppets.
Review task for children including instructions.	1b. Give PA instructions to puppets.	Instructions
Review task by having them perform. Inpress on them the difficulty of the task and thus the necessity for finding a way to make it easier.	2. Run 5-item mixed N- and -C item list. Get child to agree (as puppet) that it is a hard task.	5-item mixed list
Tell children about trick Show children how trick works -- what they are to supply.	3. Tell children you have "trick" to help them remember the two things in each picture--the "trick" is to imagine (think) what it would be like if the 1st thing did something to the second thing. For instance is we saw this [show picture of teeth and apple--stationary pictures] what could we have the teeth do to the apple? [Get response that "teeth could bite the apple"] And if we made a picture in our heads of what that looked like, it might look something like this: [show action sequence] Or we might think of something like this [show locational sequence]	Film clip of NC picture of teeth-apple Film clip of NA picture of teeth-apple Also locational picture of teeth-apple
Have children perform mnemonic in the easiest situation (with labels provided for objects and with all the cues and hints for action that physical manipulable objects present)	4. "Let's try making up stories for some more things." Have children practice making up action stories for 2 objects at a time using miniature real objects, each time having them create story and then verbalize it with a sentence [8 pairs] Name objects as present to child	2 sets of 8 pairs of objects
Demonstrate what benefit they get from using mnemonic--compared to using none (perf. on 1st list)	Run test trial	Display board Boxes

Purpose of Activity		Materials
Give children practice in working from sentence 1st; i.e., first providing sentence and then imagining what the sequence would look like (hopefully)	Have children think up action sequences for labeled pictures on TV expressing their suggestions verbally with sentences and at the same time asking them to imagine what they would look like. Show action sequence of same objects and have them label your sequence. [5 pairs]	For 5 pairs: coinc. picture of a pair followed by action sequence of same pair

Day 2 - Tutorial

Purpose of Activity		Materials
Have children perform mnemonic in hard situation (aud. mode) but with familiar pairs and ones which you can be pretty sure they can actually remember acted out action sequences. Review mnemonic	1. Review Day 1 by playing auditory list of 8 pairs learned with objects and having them remember stories and action sequences performed.	Aud. list of 8 pairs from Day 1 (ones had objects for)
Continue to show benefit for learning and keep situation similar to test situation.	Run test trial.	
Child required to make up story with objects without actually getting to manipulate them. E gets to check correctness of stories and child gets to match his internal picture of the activity with what the activity really looks like and also practices verbalizing. Child has to make verbalization 1st without getting to manipulate objects afterward--assurance child can perform verbally without object Child also gets to see that in this game there is not one right answer--but many possibilities. His answer is as good as anyone else's.	2. Guessing Game: child presented with 2 objects at a time. Child makes up "story" to himself about the two things and E tries to guess what it is (verbally) After the guess, the child shows his story to E, so she can see whether or not the guess was correct, and also then verbalizes the story. On the next item pair the roles are reversed -- E makes up the story and the child tries to guess it. [E should stress that any number of answers are correct, good ones, but she should require them to be the correct form -- the 1st Noun does something to 2nd Noun.] (8 new pairs)	Two sets of 8 new pairs of objects Display board Boxes
Show benefit of mnemonic and keep situation like test.	Run test trial with 1 object from each pair	
Children practice in different and less concrete mode - but one less remote than TV pictures.	3. Have children make up "stories" for cutout pictures pasted side by side on paper [7 new pairs] (If enough time for 2nd trial do not furnish labels on study trial - child must furnish [or remember])	2 sets of 7 cutout picture pairs - new stimuli

Purpose of Activity		Materials
Children have to provide mnemonic in a situation where manipulation and demonstration of action is impossible		
Show benefit of mnemonic and keep situation somewhat like test situation - continue to make purpose of activities clear to child	Run test trial	

Day 3 - Tutorial

Purpose of Activity		Materials
Try to push child to remember purpose, situation requiring, and nature of the sentence mnemonic - but this time providing many cues and phrasing so as to make the statement easy to make. (Do not yet require child to remember in the 1st place that he should be considering how to learn and the use of a strategy)	1. Ask children if someone asked them to remember a lot of pictures with two things in each picture, how could they remember what went with what? What trick could they use?	
Reviews mnemonic; uses "easier" situation for this 1st list of the day.	2. Review use of mnemonic by having children manipulate object pictures saw on Day 2 with stationary pictures on flannel board showing action sequences they had thought up previously. Urge rapid performance and recitation of sentences. The sequence should be:	2 sets backed pictures of 15 items which were used new on Day 2. 2 Flannel Boards
Move child toward rapid performance so he may eventually be able to perform mnemonic at the rate of the posttest (internalize the action)	(a) Child hears names of 2 objects (b) Child picks up pictures of 2 objects reciting his story as he does or shortly thereafter (c) Child performs story in brief, quickly executed motion. Test Trial with pictures	Boxes
Have child practice technique in same mode as posttest. In addition require that he also name the items himself - a more difficult task which he hopefully is now ready for. Continue to require him to provide the missing part of elaboration - in this case, the verbalization of the sentence to correspond to the action sequence (we want him to be able to perform mnemonic with varying types of input)	3. Child sees 5 pairs of stationary objects on the TV (not named) He is encouraged to name them and then construct sentences reflecting his action stories for the pairs. After each of his stories, he is shown the televised action sequence and is asked to "name" the action. Child sees 1st of each pair on test trial [no verbalization on tape]	5 new televised pairs - 1st coincidental picture of pair, then action picture of each pair, then blank space (no sound on tape)

Purpose of Activity		Materials
More difficult test trial format he must supply stimulus name himself and then remember his elaboration.	4. Child sees mixed N- and -C list on TV (10 items) Child is required to name and make up sentence out loud for each. (Rate of presentation probably slow to insure correct performance)	10 new item mixed list (N and C) on TV
Child must perform mnemonic with the 2 most difficult stimulus situations - and he must change his set for what he must provide constantly throughout the list	Test trial	
	Award Stickers - put on charts along with child's name	Stickers Charts

Day 4 - Tutorial

Purpose of Activity		Materials
One step further in phasing out specific instructions to use mnemonic - to try to get child to remember on his own to use it	1. "Today you're going to learn some more lists of 2 things that go together. What are you going to do while you watch and listen?" [If no response, say "What did you do yesterday to help yourself remember the 2 things in each picture?"]	
Specific practice in visualizing as well as providing sentences	2."Now while we make up our stories about the two things each time, we should be thinking what our stories look like. To practice thinking about what our stories look like, I'm going to try to draw your stories as you tell them to me.	2 magic slates 5 "drawable" object-pairs
Practice in verbalizing descriptive information--description must be good enough to communicate idea to another (i.e., experimenter)	1st We'll hear the names of 2 things. Then you make up a story about them. What the 1st thing does to the 2nd thing. Then you tell me what your story looks like and I'll try to draw it. Here we go. [Do for 1st 3 pairs]	
Make child do entire visualization himself.	For last two pairs - This time I'm not going to draw your story - but you think what it looks like [Perhaps draw squiggles or splotches and have child pretend they are the objects as he sees them]	
Practice doing visualizing and sentences quickly on old list	3. O.K. let's practice your trick of making up stories and thinking what they look like and see how fast we can do it [old 15-item auditory list]	Old 15-item auditory list
Practice doing both mnemonics on new list like posttest (mixed)	4. The last thing we're going to do today is a new list of 2 things. Use your trick and see how many you can remember because we're going to put it on your chart. For each one you can remember, I'll fill in a block on your chart [Mixed 10-item list]	New 10-item mixed list Charts Markers

Day 5 - Tutorial

Purpose of Activity		Materials
Fading out specific instructions - Reword yesterday's question so. less information about what task is (no mention of 2 objects)	1. "We're going to do our remembering game again today. What are you going to do to help yourselves remember?"	
Task requires performance of mnemonic without picture support. Focuses on production of sentence in clear voice. Lets child hear what he sounds like using language Give experience in benefiting from trials.	2. Play tape-recorded audio list. After each item have them record their stories. After all are recorded, rewind tape then play the tape which will have pair names, and the child's stories. Have the child listen and visualize his stories, telling you what they look like. Run test trial. 2nd trial - Have them imagine to selves.	5 pair audio list 2 tape recorders
Last step in phasing out object and picture support for sentence construction Requires child to work with internal representations of sequence. Acting out gives trainer feedback as to whether or not child is doing the task. Audio list requires concentration. Practice receiving no feedback	3. 1st 5 of 10 item audio list (ones child has seen in an action sequence before) Have child recall what he saw as he says the "story". Trainer uses some representative gesture to represent the action as child says it. Last 5 (new items easy to represent with action) Have child say sentence and represent the action himself in some way. [Do not give feedback as to correctness of responses on the test trial]	10-item audio list (5 old, 5 new the old ones having been seen in action sequences
Longer visual list practice	4. Run new 15-item visual list for charts. Give child 1 sticker to paste on his chart if he gets half the items correct, 2 stickers if he gets them all.	New 15-item visual list Charts, markers, stickers

Day 6 - Tutorial

Purpose of Activity		Materials
Last step in phasing out specific instructions Check on whether or not child performs mnemonic spontaneously	(Do not remind about mnemonic, unless child does not perform it on the first 2 items of the 1st list. If necessary, stop machine, rewind to beginning and ask child to learn the pairs as he has been practicing all week.) 1. Run 10-item auditory list for charts (5 old, 5 new) Before starting, tell the child to think hard and do it the right way.	10-item auditory list (5 old, 5 new) Charts, markers
Prepare child for situation where he must think fast to provide sentences in the time allowed. Provide incentive for speed (beat best time)	2. Have children practice saying stories as fast as they can. Use stopwatch to time and tell child how long it took after each story. Tell child to be sure to imagine visually as well. [If the child wants to operate the watch, tell him he can time the last list]	2 stop watches New 5-item audio list
Give child another opportunity to verbalize the mnemonic - but this time the child must be more complete and specific	3. Tell child that it is last time to play the games with you. Tell him you have a friend that wants to learn how to play the remembering game well and is having a hard time - and maybe the child can help [New puppet appears and asks child for help] After child explains the mnemonic, tell child to show puppet how he does tricks.	2 new puppets
Last list is as close to posttest list as possible - no feedback, 2 trials, mixed list and 20 items	4. Run 20-item mixed list - no feedback, 2 trials.	New 20-item mixed list, 2 trials
	Puppet thanks and E thanks child	

Day 1 - Didactic

Purpose of Activity		Materials
Familiarize children with trainer and situation. Review task for children including instructions Impress difficulty and need for help on them.	1. Have child try puppets on. Tell him he is going to get a chance to play the television game like he did before. Give PA instructions to puppets. Run 5-item mixed list and get child to agree (as a puppet) that it is a hard task.	2 puppets 5-item mixed list
Trainer provides elaboration - Demonstrate picture-imagined interactions in the easier situation - where objects are named for child	2. Tell child it will be easier if he watches the pictures and listens to the names. You are going to show him how to do a new trick too. This trick is - when he sees a picture of 2 things [NC picture of teeth-apple] he is to think up a picture story in his head - like this [action picture of teeth-apple] --"This 1st time, watch each picture of 2 things and I'll show you how I make up picture stories about each one. Then you close your eyes and see if you can remember how my picture story looked. Then we'll see if the trick helps you remember when you see 1 thing in each picture [5 sequences of 1st NC picture, then NA picture]	Film clip of teeth-apple in NC condition, then in NA condition 5 sequences of NC then NA pictures
Require child to produce visual elaboration on his own--but do provide labels. This task made easier by the fact that child has seen these items as action sequences on pretest.	3. This time after I show you two things (objects) you make up a picture story in your head about the 2 things [8 pairs of objects, named as presented to child. Do not allow manipulation - keep on board.] Give sticker on test trial if get at least 5 out of 8 correct.	8 pairs of objects Display Board Boxes Stickers

Day 2 - Didactic

Purpose of Activity		Materials
Review of picturing mnemonic	1. Review Day 1 by (1) presenting 4 pairs of objects first in side-by-side arrangement, then action sequence; (2) then presenting the 4 other pairs of objects and asking child to make up his own story picture	8 pair list - Manipulable objects
Introducing sentence-Mnemonic - trainer demonstrates (in easier situation where he sees stimuli and has them labeled for him)	2. Tell child that today when he sees and hears the 2 things you are going to tell him a talking story about them and he is to repeat it. (1) Present each pair of objects separately on a board naming each. (2) Read sentence (3) Ask child to repeat sentence (Do not manip. objects or allow child to) (4) Present 1st of each pair singly on board, then return to box	Display board 2 sets of 8 pairs objects Boxes
Child practices mnemonic on own - but in the "easier situation" as above	3. Tell child that this time he is to make up his own talking story about the 2 things he is going to see on the papers. (as show pictures, label them)	7 pairs cutout pictures - new
Child "eased" into harder situation where objects not labeled and he must think of them himself	Run test trial (If time for 2nd trial, don't label during study trial) [Do not give child feedback about type of sentence he forms or way he words it]	

Day 3 - Didactic

Purpose of Activity		Materials
Review of sentence mnemonic in new medium	1. Review Day 2 by (1) Reading sentences for the 5 pairs of pictures presented on the flannel board.(2) Have child repeat sentences. (3) For the next 10 pairs, have the child provide the sentences (4) Run test trial	Flannel Board 15 pairs of cutouts Boxes
Review picture mnemonic by demonstration again. This time list is mixed - so only have still pictures sometimes as stimuli	2. Tell children you are going to have them do the trick they did the other day - the "picture stories". Show him a 5-item mixed list on the TV - not labeled on TV but by you. For each pair show him action sequences, then a blank so he can picture the action he has just seen	5-item televised mixed list followed by action picture then blanks
Child tries sentence mnemonic on mixed list	3. Show child a mixed list on the TV and tell him to think up some action picture stories for them	10-item mixed list
	Run test trial	
Incentive	Award stickers for correct answers [at least half right?] Put stickers on charts	Stickers Charts

Day 4 - Didactic

Purpose of Activity		Materials
Review of picture mnemonic	1. Remember yesterday we were practicing making picture stories in our heads. For these 1st ones I'll tell you about the two things, then I'll show you what the picture in *my head* looks like on this chalkboard	
Demonstration	*3 pairs* - For each one (a) read names (b) moment of silence when you and child think of story pictures in your heads (c) draw interactive picture on chalkboard on slate	2 Magic Slates 5 pairs of names of objects that are drawable
Require child to perform	*Last 2 pairs* - (a) read names (b) instruct child to make picture story in head (c) Run test trial	
Sentence mnemonic in more difficult situation - only hear names instead of seeing pictures or objects Demonstration	2. Today we're going to practice our other trick again - where we make up *talking* stories. But this time we're just going to *hear* the names of the 2 things - we won't *see* any pictures. This first time I'll say the talking stories after we hear the names of the 2 things and you repeat the stories after me (say what I say) [Old 15-item list do rapidly].	Old 15-item auditory list
Child performs with mixed list which sometimes gives picture support but requires him to shift from mode to mode	3. O.K. This time *you* make up the talking stories. Sometimes you'll see the 2 things, sometimes you'll hear about them but each time make up a talking story like I did.	New 10-item mixed list
	We're going to use your charts to see how many you get right - so think hard and make up some good stories	Charts, Markers

Day 5 - Didactic

Purpose of Activity		Materials
Review Added interest of hearing child's voice providing sentences	1. "Yesterday we practiced making up talking stories. Let's see if you remember how we do it." "First I will play some stories that other children have made up for the 2 things, then you'll get to record your own for another child to hear. First listen and repeat this little (girl, boy)'s talking stories." For 1st 3 pairs (a) Play pair (in E's voice) (b) Play "story" (in child's voice) (c) Have child repeat the "story"	4 tape recorders Tape of E reading 3 pairs of words followed each time by child's voice saying sentences
Child performs task	For next 2 pairs (a) Play pair (in E's voice) (b) record child's story on different tape recorder Run test trial on the 5 pairs Rerun tape with the child's voice on it and another test trial.	Tape of E reading 2 word pairs followed each time by a blank.
Put 2 mnemonics together for the 1st time.	2. "Now we're going to put our two tricks together. We're going to make picture stories and talking stories at the same time. This first time I'll do it for you to show you how.	10-item auditory list (5 for which you also have televised action sequences)
Demonstration with visual support and next with less visual support	Run 10-item audio list - 1st 5 items (a) Run item pair, (b) Give sentence, say "and if I made that into a picture story it would look like this in my head.[show TV action], (c) have child repeat and imagine sequence after disappeared. 2nd 5 - (a) make up sentence and partially act out and describe action sequence using hand and facial expressions as say sentence. (b) Have child repeat sentence and imagine what you are describing. (c) Run test trial - no feedback	
Child performs 2 tasks on his own - No feedback on performance of mnemonic but does get on learning performance.	3. Tell child this time he sould make up talking stories and at the same time think what his story looks like - make up action picture story in his head and show you about them. (new 15-item visual list) Record on charts, award stickers	15-item new visual list (-C condition) Charts, Markers, Stickers

Day 6 - Didactic

Purpose of Activity		Materials
Motivational gimmick	1. Get out stopwatch - show to child and tell him you want him to help you find our how long you spend working today. Show child how watch works and have him start it (while you hold it) then put away.	2 stopwatches
More practice with using both mnemonics - 1st demonstration without any picture support	2. Tell child that he is going to practice using the 2 tricks together like yesterday - the talking stories and picture stories both. "This first time listen and I'll tell you what they look like." For each of 5 items (aud.) (a) Run auditory taped pair (b) Provide sentence (c) Have child repeat sentence (d) Describe how it looks in your head. Run test trial	New 5-item auditory list
Performance in hardest situation Chart for motivation and feedback	3. Tell child to perform these operations himself on the 10-item list that he will hear next. On last trial record the number of correct responses on his chart.	10-item auditory list (5 old, 5 new)
Last list is as close to posttest as possible.	4. Tell child that a new puppet wants to come and watch him do the last list and see how much he has learned - so do the best for him [puppet]. Puppet appears and is introduced Tell child to do his 2 tricks - talking stories and picture stories - for himself. [20-item mixed list - 2 trials, no feedback]	New 20-item mixed list - 2 trials Puppet

Day 1 - Practice

Purpose of Activity		Materials
(Same as other treatments)	1. Have children try puppets on and tell them the puppets are going to get a chance to play the television game like the children did before. Give PA instr. to puppets. Run 5-item mixed N and C item list and get child to agree that it is hard task	5-item mixed N and C list. 2 Puppets
Try to get child to discover rehearsal "trick" on his own Let child practice rehearsal with list- so he can see benefit.	2. "Let's see if we can find a way to help you remember them better. Tell child to listen carefully while you read a pair of words and see if he can remember them for 10 seconds. (Show him stopwatch) Read 2 words ______ ______ "Now can you remember these two words 20 seconds?" (Read 2 different words ___ ___) If child was repeating words to himself during the 20 seconds ask him what trick he used. If no response - "Were you saying the 2 words to yourself while we waited? Well, (whether or not he suggests or admits it) let's use that trick to try to remember some more sets of 2 things. You say the names to yourself after the voice does. I'll show you 2 things and tell you what they are and you say the names after me. (Run list of 8 pairs of objects) Present side-by-side and do not allow manipulation by child. This second time say the names twice after I tell you.	2 Stopwatches 2 pairs of words (not recorded) 8 pairs of objects Display Board Boxes
Give child chance to practice rehearsal in different medium.	3. Let's try our trick with ones on the TV - 5 NC pairs. Run twice - 2nd time give sticker if get them all.	5 televised NC pairs Stickers, Charts, Markers

Day 2 - Practice

Purpose of Activity		Materials
Practice with familiar list (easier task 1st)	1. Review Day 1 by playing auditory list of 8 pairs learned with objects yesterday. Remind him to rehearse each after hearing it. Run test trial	Old 8-item auditory list (ones have objects for)
Help him understand that groupings are arbitrary Practice Puppet is motivation Child gets to practice learning in different role - still must remember pairs.	2. Tell child he is going to get to make up his own pairs of 2 things for you and a puppet to learn. Display all the objects on a board (naming them as you point to them) and ask him to group them into pairs. Then have him name both members of each pair for you and a puppet he holds. Alternately you and the "puppet" rehearse items presented. Then tell him to put one from each pair in the box and then name the one remaining and you'll try to guess what one is missing from each pair. Have the puppet (therefore the child) guess every other one. Have child say if you or the puppet are correct.	2 sets of 8 pairs of objects Display Board Boxes
Different medium more practice	3. Tell him you have one for him to learn now, using his rehearsing trick. [7 cut-out picture list - 2 trials if time]	7 pairs cut-out pictures - new
Try to insure you will get him back - carrot in front of nose	Tell him on another day he'll get to record a list on the tape-recorder	

Day 3 - Practice

Purpose of Activity		Materials
Whet interest	1. Tell children "later we're going to have a puppet contest, but 1st let's do one of our learning games to see if you remember how to do the trick about saying the names (words) again after you hear them.	
Incentive	I'll give you another sticker for your chart if you get most of them. (Use 15-item list on flannel board) As you say pair, put them side-by-side on flannel	Flannel Board Cut-outs - 15 pairs
Child gets to participate. Practice	board, have child rehearse, then have him take them down giving you one and putting the other in a box. Run test trial with ones in your hand.	Stickers Charts Boxes
Incentive	Award sticker if performance good (i.e., over half right)	
	2. Puppet Contest - Have child select	2 Puppets
Low key competition Child against himself really but make him conscious of performance	2 puppets he wants to compete in contest - tells puppet how to do task. [1st puppet learns 1st list and 2nd puppet , the 2nd list - the child will be answering for both]	(?mixed list)
	(a) 1st list - 5 "-C" items on TV - 2 trials Fill in blocks on "puppets" chart.	5-item visual list, no labels 2 trials
	(b) 2nd list - have child put on other puppet and do 1 trial of 10-item mixed N + C item list. At test trial fill in 2nd puppet's chart.	10-item mixed list 1 trial televised
	(c) Have child compare charts of 2 puppets and declare the winner. Shake hands and clap for winning puppet.	Puppet charts Markers
	Tell child he will get to fill in some blanks in his own chart tomorrow.	

Day 4 - Practice

Purpose of Activity		Materials
Different medium	1. Tell child you are going to draw some pictures for him to remember today. Draw 5 pairs in side-by-side arrangement. Label as you draw them. Have auditory test trial	2 Magic Slates 5 pair list of "drawable objects"
Practice learning at faster pace	2. Do old 15-item list auditorially to practice doing rapidly - at pace. Tell children are practicing for tomorrow when will time how long it takes to learn list.	old 15-item auditory list - do at pace
Mixed list practice Motivation	3. Do new 10-item list "for charts" - mixed N + C conditions [Do marking for child]	10-item mixed list televised
	Point to one on charts will do tomorrow	Charts Markers

Day 5 - Practice

Purpose of Activity		Materials
Gimmick - different way to practice	1a. Display stickers - have children stick them together then record them as PA list. [You record test trial] 1b. Play back and have child as puppet use rehearsal and respond	5 pairs [stickers available for items] 2 tape recorders
Practice having no feedback	2. Run 10-item list as "warmup" for one do for charts. No info. on results [5 old, 5 new audio.]	10-item audio (5 old, 5 new)
	3. Show child chart and the blocks he will fill in	15-item new visual list
Longer, new list Motivation	Run 15-item new visual list for charts. Give 1 sticker if get half right, 2 if get all.	Charts Markers Stickers

Day 6 - Practice

Purpose of Activity		Materials
Gimmick to keep practice interesting	1a. Show child stopwatch and how it works and say you want him to time how long it takes to do the various lists. Let him push button to start watch - but you hold watch and keep out of sight until the end of the list.	2 stopwatches
	1b. Run 10-item auditory list for the charts (5 old, 5 new) Mark charts. Bring out watch and stop it. Look at time with child, Mark down time so can compare.	10-item auditory list (5 old, 5 new) Charts, Markers
Incentive for working at faster rate	2. Say you will run the next list a little faster so "we can beat our record" - so think fast. [New 5-item auditory list - 2 trials] Check time and mark down.	New 5-item auditory list, 2 trials
Final practice at "post-test-like" list	3. Say to child that they can see if this practice has helped the child to be able to do a long list. "Try hard and see how many you can get." (20-item mixed list - no feedback - 2 trials)	New 20-item mixed list - 2 trials

Appendix B

List of Target Words for Vocabulary Training

Words taken from Peabody Picture Vocabulary Test (Form B)

ambulance
argument
balancing
bannister
barber
binoculars
captain
cash
climbing*
cobweb
cone
engineer
globe
gnawing
hive
hook
hydrant
idol
insect
locomotive
parachute
peeking
pulling
rat
reel
river
saddle
sail
temperature
time
trunk
tweezers
walrus
wasp
weapon
whale

Words suggested by the Picture Interview

apron
branches
clothes
clouds
desert
donkey
drill
elephant
giraffe
guitar
hammock
hanging
happen
holding
jungle
knight
lion
ocean
pliers
sand
scared
screwdriver
slide
stool
stilts
stripe
swinging
talking
vase
waves
workbench
yelling
(basic color terms)

Additional Words from the PPVT

(The following words were taught to some of the kindergarten subjects when they had mastered the above words and there was time left over during the training sessions. They are items 61-80 of Form B.)

assistance
astonishment
autumn
casserole
chef
cobbler
construction
dissatisfaction
erecting
filing
harvesting
horror
oasis
observatory
ornament
scholar
shears
soldering
thoroughbred
tread

---

*Actually suggested by PI, but happened to be in PPVT.

Appendix C

List of Elaborative Constructions used for Sentence Training

Elaborative elements: adjective, possessive noun, prepositional phrase, and relative clause (containing no object) attached to the underlined constituents in the following sentence patterns:

1. Subj. + be + *Noun*
2. *Subj.* + be + Adj.
3. Subj. + be + *Prep. Phr.*
4. *Subj.* + Verb + Reflexive Prn.
5. Subj. + Verb + *Obj.*
6. Subj. + Verb + Poss. Prn. + *Obj.*
7. Subj. + Verb + Indir. Obj. + *Dir. Obj.*
8. Obj. + be + Verb + by + *Subj.*

Indirect discourse with the above patterns as the embedded sentence in the following constructions:

1. Subj. + say + that + Sentence
2. Subj. + tell + Obj. + that + Sentence

Adverbial clauses introduced by the conjunctions *because*, *when*, *before*, and *after*.

Relative clauses containing objects and modifying either subject or object of the matrix sentence.

Appendix D

Sample Lesson Plans for Language Training Sessions

Vocabulary Lesson # 4

A Trip to the Toy Store

Target Words: cash, cone, ambulance, parachute, saddle, balancing (review) locomotive, knight, guitar, binoculars, donkey (review)
Optional Words: soldering, ornament, thoroughbred, shears, treat

We're going to take a pretend trip to the toy store. How will we get there? Let's take a ____ (show picture of bus). What color is the bus?

We'll have to pay the bus driver for taking us. Let's get out some cash (show both coins and bills) This dollar is cash and these pennies and nickels are cash. All money is called cash. The dollar and the coins are both what? _____.
Let's give the driver the dollar and see if he gives us back any cash. (Driver gives back 2 coins) Did he? What did he give back? (If child says money or pennies at any time, ask "what's the other word for both pennies and dollars?")

OK, we're there. Oh look. There's a man selling ice cream. We can either buy it in a dish or on a cone. Which are you going to buy - a cone or a dish. (Ask each child in turn) But, we'll have to give the man some cash to pay for it. Johnnie you give him some cash and he'll give you what you asked for (cone or dish of ice cream) (Repeat for each child) (Give each child the money to pay)

Oh, here's the toy store. Let's go in. I wonder what he has - maybe he has some toys on sale today.

Here's the first shelf of toys. (soldier wearing parachute, policeman, knight)
(a) Ask the children to name toys they see; if they can name them, add the (workman with his soldering kit); if they have trouble naming, name for them, have them repeat the names, and give a little description of them.

(b) Ask children what you should buy if:
1) You want to play that you are a princess and a dragon is after you and you need someone to save you and fight the dragon.
2) You are going to jump out of a plane and you need someone to show you how to do it so you won't get hurt.
opt. 3) Your water pipes at home are broken and you need someone to come and fix them so the pipes stay together.

(c) Ask children to pick out a toy they would like to buy and give the store owner some cash. (Give each child some money, have him name the item he wishes to buy and then "pay" for it)

Repeat for each shelf of toys

Shelf 2: seal balancing ball, horse wearing saddle, donkey, (Optional toy - thoroughbred horse)

Questions: What should you buy if:
1) You want to pull some long ears.
2) You'd like to watch someone make a ball stay on his nose.
3) You'd like to sit on a seat made to go on an animal.
opt. 4) You'd like an extra special horse.

Shelf 3: ambulance, truck, car, locomotive [optional - tire with good (rough not smooth) tread]

Questions: What should I buy if:
1) I want something that will pull cars along a track.
2) I want to play doctor and I have a sick patient that needs to go to the hospital.
opt. 3) I need a very good new tire for my car.

Shelf 4: guitar, banjo, binoculors, sun glasses (optional - hair ornaments)

Questions: What should I buy if:
1) I want to see something a long way away.
2) I want to play a song on an instrument that is not round.
opt. 3) I want something pretty to put in my hair.

Vocabulary Lesson # 9

Detectives in the Attic

New Words: scared, bannister, insect, clothes, vase, attic, wasp
Optional: horror

Review: trunk, weapons, rat, gnawing, guitar, hook, saddle, cash, hydrant, binoculars, locomotive, globe, apron, peeking, cobweb, hive

1. Have you ever been scared before? What has scared you? (Each child says one thing that has scared him) If children do not know meaning of word, give instance of something that happened to you and show by facial expressions and tone of voice. (Use horror instead of scared for better groups)

2. I know a boy and a girl who became scared when they went up in their attic. (Explain that attic is place on top of house where keep things in boxes and barrels and piles) Show picture of attic.

   Let's be detectives and find out what scared the boy and girl. (What filled them with horror - optional phrase)

3. First the boy and girl climbed the stairs and held onto the what? (bannister) Show picture. They held on so they wouldn't fall. What did they hold onto?

   When they got into the attic they saw many things. Let's see if you can tell which thing scared them.

4. First they saw some insects - a what (bee) The bee was flying to his home which is called a what? (hive). Another insect they saw was a what (spider) and he was spinning a what (cobweb). Then they saw another insect - a wasp (What will a wasp do if you're not careful?) A bee is an insect, a wasp is an insect, and a spider is a what? (insect) These bugs are all what? (insects)

   Do you think it was the insects that scared the boy and girl (No)

5. Then they saw a big old what? (trunk) Do you think the trunk scared them? In the trunk they found some old what? (clothes) Do you think the old clothes scared them?
   Farther down in the trunk they found a pretty what? (vase) Did scare?
   Next they found some old what? (weapons) Did these scare?

6. As they were closing the trunk they saw a what? (rat) and the rat was what? (gnawing) on some wood. Did the rat scare them?

7. Then the boy and girl saw a what in the corner? (guitar) Did the guitar scare them, do you think?

8. Then - all of a sudden - they saw something large move in the corner. They looked closer and saw a very strange looking animal doing what (peeking at them) They looked at each other and then back at the animal and all of a sudden it said "Boo" They were so scared, they turned around and ran down the stairs, holding tightly to the what? (bannister) and the animal hid himself behind something in the attic.

Let's see if we can find him.

Let's look under the,e.g., bannister. Who can find the bannister (Have child point)

OK Johnny, look and see what is under the bannister (Have child say what he sees)

Sentence Construction Lesson # 3

Subj. + be + Prep. Phrase and Review of Lessons 1 & 2

Materials to hide:
same toys as in 1st lesson
pencil
toy horse
toy comb

Materials to hide in, behind and under:
red box and green box
2 med. size "brown" boxes
one paper bag
teacher's sweater (or coat) and one of the children's sweater's or coat (same item in any case).
2 animated toys, same object, different actions, to hide behind.

Activity 1

"Did you ever forget where you put something, and then you just couldn't find it? Are you good at finding things when your mother forgets where she puts them? Let's see if you are good at remembering where things are. I'm going to hide some things and you try to remember where I put them."

For a slower group you may wish to hide only 5 + 4 objects at a time. For a sharp group hiding all of the items at once may make the task more interesting and challenging.

Hide items so that they are not visible but children watch you hide them. Ask Fletcher (the puppet) the first one so that the model is provided. If time is going well, replace each item as it is guessed correctly so that you can ask another child, thus giving each child a turn with each model. At any rate do this with the last pattern, S + be + prep. phrase + rel. As a child "guesses," in correct form only, show him whether he is right or wrong.

The _______ is in the sack.

The _______ is in the red box.
The _______ is in the green box.

The _______ is behind the box on the floor.
The _______ is behind the box on the chair.

The _______ is under the teacher's coat (sweater).
The _______ is under Mary's coat (sweater).

The _______ is behind the _______ that _______.
The _______ is behind the _______ that _______.
(insert name of animated toy + verb)

To elicit guesses, ask "Where is the _______."

Activity 2 Review of lessons 1, 2, & 3

Sing the refrain to the tune of "I'm a little teapot", pause and ask the question, furnishing the first part and getting a completion response from the child, sing the refrain again, encouraging children to join you, ask the next question, etc.

Refrain: Fletcher is a kitty,
Fletcher is black,
Fletcher is on the table, (demonstrate)
Fletcher is on the sack. (demonstrate)

1. "Tell us what he does?"

Fletcher is a kitty (that) (drinks) . demonstrate w/ paper cup.

Sing refrain.

2. Bring out paper cat. "We now have 2 kitties. (Put Fletcher on table, paper cat on floor.) "Tell us where the kitty that is Fletcher is."

Fletcher is the kitty on the table.

Sing refrain.

3. "What box is Fletcher on?"

Fletcher is on the box under the table .
(switch) Fletcher is on the box on the chair .

Sing refrain.

4. "What (animated toy) is Fletcher behind?"

Fletcher is behind the ________ that ________.
(switch) Fletcher is behind the ________ that ________.

Sing refrain.

5. Present another paper cat and announce that his name is also Fletcher. "One Fletcher is cloth and one Fletcher is paper - one is black and one is pink. Which Fletcher is cloth? ... Which Fletcher is paper?"

The Fletcher that is black is cloth .
The Fletcher that is pink is paper .

6. "One Fletcher is on the table and one Fletcher is where? One is black and one is pink. Which Fletcher is black? Which Fletcher is pink?

The Fletcher on the table is black .
The Fletcher under the table is pink .

Refrain.

Sentence Construction Lesson # 9

Indirect Discourse

Theme: Secrets

Activity 1

Fletcher is going to tell secrets today. He is going to tell everyone a secret. Some of Fletcher's secrets are true, and some of them are silly. If Fletcher tells you a true secret, you pat him on the head. If Fletcher tells you a silly secret, you shake your finger at him. Then you tell us what Fletcher said.

Model Sentences: (For better groups use as much complexity as desired. For slower groups, omit complexity).

Fletcher said that:

1. Mary is on a chair (next to Johnnie).
   A (green) rabbit is under the (teacher's) chair.
2. Mary is a girl (who is sitting) (on a chair).
3. Johnny is wearing a (blue) shirt.
   Johnny is wearing a (big) flower (on his head).
4. The (teacher's) pencil (or on the table) is yellow.
   The teacher ('s hair) is green.
5. Johnny gave Mary a horse (which talks).
   Johnny gave a clown (a flower).
   (a dog) (which talks).
6. Johnny was chased by a (big) rabbit.
   Johnny was hit by a clown (who laughed).

Activity 2

Now we are going to play a trick on Fletcher. He is going to tell (the pink rabbit) (you) a secret. (If use rabbits, add that rabbit will tell a child the secret). You promise to keep the secret, but then you tell us what it is.

Use rabbits for Fletcher to tell secrets to for sharper groups, and add in complexity as desired. For slower groups do not use rabbits or complexity.

Thus children will respond with: Fletcher told the pink rabbit that....
or just Fletcher told me that....

Sentences:

1. Johnny is a (nice) boy.
   Mary is a (pretty) girl (who is sitting).
2. (The teacher's) Some feet are on the floor.
   The (teacher's) (brown) purse is on the table.
3. Johnnie is wearing (brown) shoes (on his feet).
   Mary is wearing a (red) ribbon /bow/ (in her hair).

4. Some (or the teacher's) shoes are white.
   The (teacher's) pencil is yellow.
5. Fletcher told the children (silly) secrets.
6. The purse (on the table) was brought by the teacher.
   A (yellow) pencil is being held by the teacher.

Note: For neither activity is it necessary to use all of the sentences. This should depend on your time. But try to use one sentence from each type.

Appendix E

Instructions for Transcribing from Tapes

I. Procedure

- A. Use yellow paper for rough draft and white paper for final copy.
- B. Listen to the tape recording once completely, typing rough draft for general idea of transcript.
- C. Listen to the tape as many times as necessary, writing in additions and corrections with black pencil.
- D. Type final draft.
- E. Double-check transcript by listening to the tape on a different day.
- F. When the child's speech seems especially clear, type a final copy of transcript only and write in additions and corrections with black pencil.

II. Margins

- A. Use wide margins as follows:
  2 inches from top on first page, 1-1/2 inches from bottom;
  1-1/2 inches from top and bottom of remaining pages;
  left margin - 20
  right margin - 65
- B. Indent responses of S five spaces

III. Headings

- A. Head the first page of a transcript as follows:
  last name of S, first name
  side number of tape, counter number
  school
  date
  last name of E
- B. Head remaining pages as follows:
  page number, centered
  last name of S, first name
  counter number
- C. Leave 4 spaces before each picture is introduced and indicate picture number in parentheses as E introduces each picture.
  Example: (picture 1)

IV. Spacing

- A. Single space remarks by E.
- B. Double space remarks by S.

V. Punctuation

- A. Do not use commas, colons, semi-colons, or dashes within an utterance.
- B. Use question marks where appropriate at the end of an utterance.
- C. Enclose in parentheses all comments of the transcriber, alternate interpretations, etc.
- D. Indicate pauses by one asterisk for a short pause and two asterisks for a long pause.

E. Do not break up words at the end of a line.
F. Use question mark within parentheses to indicate uncertainty about interpretation of a word or phrase.
G. Capitalize the first letter of the first word of all responses by S. Use a period after every response by S.

V. Spelling

A. Use standard way of spelling consonant sounds.
B. Use standard way of spelling vowel sounds--unless you can't tell what the word is.
C. Indicate phonemic spelling with slashes when the sounds are clear even though the word(s) cannot be determined.

Vowels:

a	father
ae	bat
aw	hawk
ey	bait
iy	feet
i	bit
ay	bite
oh	boat
oo	boot
oi	oil
ou	bout
uh	butter

D. Spell contractions in the standard way when they occur. But in those ambiguous cases which might be either a contraction or a plural noun, use -s, as in boy-s.
E. Write "because," not "'cause," and "going to," not "gonna." Spell out in standard way most deletions and slurrings. Spell out stutter sounds of a word only if clear.
Example: Fi-five.

VI. Miscellaneous

A. Write "interrupted by...") in parentheses to indicate interruptions by S or E or other.
B. When uncertain about a word or words, use a question mark in parentheses, followed by x's to indicate the number of syllables that cannot be determined, and give the number of unknown syllables.
Example: (? xx 2 syllables)
C. For picture interview, put tape recorder counter number at the beginning of each page of transcript. For sentence imitation, put counter number every five sentences.

Appendix F

Scoring System for the Sentence Imitation Test

The present system was devised for the initial analysis of SIT protocols. (The rules for transcribing SIT tapes should be consulted before this system is used.) Each response from the SIT is assigned to one and only one of the categories described below, according to the kinds of "error" which the response contains. The following general rules and guidelines apply to the scoring of all responses.

General Rules and Guidelines

1. Only the subject's (S's) response to the first presentation of each item is scored, unless there is a note on the protocol to the effect that there was some interruption during the first presentation or response.
2. Extraneous comments by S before, during, or after a response should be ignored.
3. False starts and repetitions by S within a response should be ignored (but see category 4m). If S gives correct form first and then changes it, score the final version.
4. Ignore pure articulatory or phonological errors, such as wunch for lunch, or eatin' for eating. (These errors are usually not indicated on the protocol anyway.)
5. The categories are cumulative in the sense that a category 3 response may also contain category 2 errors, a category 4 may also contain category 2 and 3 errors, and so on. A response is always assigned to the highest numbered category possible, given the errors which it contains. However, when an error can be placed in more than one category, give S the benefit of the doubt and assign the lowest numbered category. By way of illustration, the error below could be an omitted copula (4) or a change in number, with contraction (3), so it is scored as a 3.

   MODEL: THE NEW TEACHERS ARE PRETTY
   response: the new teacher-s pretty
6. In the few cases where E has read an item incorrectly, score the response for its conformity to the sentence which actually was presented.
7. Listen again to those responses for which sound cues might be critical in scoring.

Scoring Categories

1. - Verbatim
2. - Small Optional Changes (Little or no change in meaning)
   a. Change of article where optional.

      A MAN
      the man

b. Optional addition or deletion of relative pronoun.

A BOOK THE TEACHER WANTED
a book that the teacher wanted

THE MAN WHO THE CHILDREN SAW
the man the children saw

c. That substituted for who in relative clause.

THE GIRL WHO FELL
the girl that fell

d. Who substituted for that in relative clause after animate nouns.

DOLL THAT WALKS	CAT THAT ATE
doll who walks	cat who ate

e. Deletion of optional that as sentence introducer.

KNOWS THAT THE BOYS ARE SICK
knows the boys are sick

f. Contractions.

A CLOWN IS A FUNNY MAN	BECAUSE	UNTIL
a clown-s a funny man	'cause	'til

g. Sentence begun with and.

h. Pronoun change compatible with model.

BIRD IS BUILDING ITS NEST	THE CAT IS WASHING ITSELF
bird is building her nest his	the cat is washing himself herself

3. - Minor Syntactic and Lexical Changes Maintaining Standard English
(More noticeable change in meaning)

a. Change in number, tense, or aspect.

BOY	HORSES EAT	LADY LIKES
boys	horse eats	lady liked

DOCTOR IS	CHILDREN PLAYED
doctor was	children were playing

b. Possessive pronoun replaced by or replacing article, when possession is obvious.

BIRD IS BUILDING ITS NEST	CHILDREN TOLD THE TEACHER
bird is building the (a) nest	children told their teacher

c. Addition of -y to certain animate nouns.

HORSE	DOLL	KITTEN
horsey	dolly	kitty

d. Substitution of words closely related in meaning and/or sound.

CHILDREN	MOTHER	SLEEPING
kids	mama	sleepy
DOLL	SINGS TO HERSELF	BOYS ARE SICK
dog	thinks to herself	boys are six

WERE	WHEN	CALLED	WHEN	BECAUSE
weren't	while	called on	because	when

e. Substitution of the for some or vice versa.

f. Deletion of article, some, or possessive pronoun where possession is obvious before all plural or mass nouns, and before unmodified singular mother, teacher, doctor, baby, and kitty, and school.

g. Addition of elements which have little or no effect on meaning.

THE DOLLS ARE ON A CHAIR BY THE DOOR
the dolls are over on a chair by the door

THE CHILDREN TOLD THE TEACHER THAT SOME BIRDS WERE SINGING
the children went and told the teacher that some birds were singing

h. Addition of redundant little.

THE BABY
the little baby

THE KITTEN
the little kitten

But score as 6: THE GIRL
the little girl

4. - Minor Syntactic and Lexical Changes Resulting in Nonstandard English (includes dialect and immature forms)

a. Lack of agreement between subject and verb, as in teacher know; children likes; rabbits was caught.

b. Lack of agreement between pronoun and antecedent, as in boy made some boats after they heard a story.

THE CHILDREN LIKE THEIR TEACHER
the children like his teacher

But score as 6: The girl's father cut herself

c. Omission of possessive inflection, copula, or auxiliary.

THE GIRL'S BROTHER
the girl brother

HAT IS RED
hat red

KITTEN IS EATING
kitten eating

THE CHILDREN WHO RAN WERE HOT
the children who ran hot

d. Nonstandard use of a, as in a old man; a children.

e. Nonstandard form of relative pronoun, as in girl what fell; apples who fell; doll tha(t)s walks.

f. Nonstandard forms of verbs, nouns, and pronouns, as in ranned; peoples; childrens; theirselves; themself; hisself; her own self.

HIS LUNCH
he lunch

g. Pronominal apposition.

THE GIRL'S FATHER CUT HIMSELF
the girl's father he cut himself

h. Nonstandard substitution of words closely related in meaning and/or sound.

THEY HEARD A STORY
they listened a story

THE GIRL WHO FELL
the girl who felt

i. Nonstandard verb phrases

THE GIRL FELL
the girl was fell

THE GIRL'S FATHER CUT HIMSELF
the girl's father to cut himself

But score as 6 when to means in order to.

j. Substitution of this for a, the, or some.

k. Use of some before singular count nouns, as in some man.

l. Nonstandard sequence of tenses.

THE BOYS MADE SOME BOATS AFTER THEY HEARD A STORY
the boys make some boats after they heard a story

But score as 3: THE CLOWN THE PEOPLE WATCHED WAS FEEDING A RABBIT
the clown the people watch was feeding a rabbit

m. Reinstatement of verb (and pronoun appositive to indirect object) between indirect and direct objects. (Must be preceded by pause.)

THE MOTHER MADE THE GIRL WHO LAUGHED SOME COOKIES
the mother made the girl who laughed * made (her) some cookies

n. Pronoun reinstating subject after verb (if pronoun is understood to be reflexive).

THE BOYS MADE SOME BOATS
the boys made them some boats

o. Omission of article before singular count nouns except for unmodified mother, teacher, doctor, baby, or kitty.

THE BOY IS EATING
boy is eating

THE LADY WAS A TEACHER AT THE SCHOOL
the lady was teacher at the school

5. - Agnate Sentences and Other Responses with Base Structures Similar to the Model (but score as 6 approximate agnates which add a nuance of meaning that is not implied by the model, e.g., THE GIRL WHO OPENED THE DOOR SEES AN AIRPLANE, the girl opened the door to see an airplane). (Includes clauses which cannot stand alone)

a. Transposition of verb phrases, clauses, or parts of clauses.

THE GIRL WHO FELL IS CRYING
the girl who is crying fell

THE HORSES EAT THE APPLES THAT THE BOY BRINGS
the boy brings the apples that the horses eat

THE BOYS MADE SOME BOATS AFTER THEY HEARD A STORY
after they heard a story the boys made some boats

b. Transformation of relative and adverbial clauses to coordinates. (Conjunction and/or personal pronoun may be deleted and nonstandard verb construction may be used.)

THE GIRL WHO FELL WAS CRYING
the girl fell (and) (she) was crying

THE GIRL HIT THE BOY WHEN HE TOOK THE CANDY
the girl hit the boy (and) he took the candy
when the girl hit the boy and he took the candy

But score as 6: the girl hit the boy (and) took the candy

THE DOG THAT CAUGHT THE CAT IS DRINKING SOME WATER
the dog caught the cat and (he) (is) drinking some water
the dog caught the cat * (is) drinking some water
(Score as 5 _only_ if pause after _cat_—or some other prosodic cue—indicates coordinate verb phrases. Otherwise, score as 6)

THE WINDOWS WERE BROKEN BY THE BOY WHO RAN
the windows were broken by the boy (and) he ran
But score as 6: the windows were broken by the boy (and) (was) ran

THE CHILDREN WHO RAN WERE HOT
the children (were) ran and (were) hot
the children (were) ran (and) were hot
But score as 6: the children (were) ran hot

c. Transformation of relatives and adverbials to participial phrases.

THE LADY LIKES THE GIRL WHO IS WATCHING THE BABY
the lady likes the girl watching the baby

THE GIRL HIT THE BOY WHEN HE TOOK THE CANDY
the girl hit the boy taking the candy

d. Transformation of relatives to adverbials and vice versa.

THE GIRL WHO FELL WAS CRYING
the girl was crying because she fell

THE GIRL HIT THE BOY WHEN HE TOOK THE CANDY
the girl hit the boy who took the candy

e. Transformation of passive to active (or vice versa).

THE DOCTOR WAS CALLED BY AN OLD MAN
an old man called the doctor

f. Use of _to_ with indirect object, following direct object.

A MAN GAVE THE GIRL'S DOG SOME BREAD
a man gave some bread to the girl's dog

g. Negative of model sentence (and not just a sound-alike).

A MAN SAW THE CAR BY THE HOUSE
a man didn't see the car by the house

h. Reinstatement of noun where model uses pronoun.

THE GIRL HIT THE BOY WHEN HE TOOK THE CANDY
the girl hit the boy when the boy took the candy

i. Nonstandard form of indirect discourse.

THE TEACHER KNOWS THAT THE BOYS ARE SICK
the teacher knows are the boys (are) sick

j. Transformation of model to yes/no question.

THE BOY IS THE GIRL'S BROTHER
is the boy the girl's brother?

6. - Failure to Maintain Approximate Base Structure of the Model (Major change or loss of meaning)

Any response which cannot be scored in the first five categories should be assigned to category 6. The following subcategories illustrate some of the more frequent errors in category 6. They may also clarify the preceding categories by providing contrasts.

a. Omission of major element in sentence.

A MAN GAVE THE GIRL'S DOG SOME BREAD
a man gave the dog some bread

THE PICTURE IN THE BOOK WAS FUNNY
the picture was funny
the picture was in the book

AFTER THEY HEARD A STORY
after a story

THE TEACHER CALLED THEM
the teacher called

b. Rearrangement of sentence elements into different base relations.

THE SHOES WERE IN A BROWN BOX
the brown shoes were in a box

THE MAN WHO THE CHILDREN SAW WAS DRAWING A PICTURE
the man who saw the children was drawing a picture

THE DOG THAT CAUGHT THE CAT IS DRINKING SOME WATER
the dog caught the cat drinking some water
(Score as 6 when no pause after cat or other cue to indicate that dog is subject of drinking.)

But score as 5: THE PICTURE IN THE BOOK WAS FUNNY
the funny picture was in the book

c. Use of personal pronoun which changes base structure of model.

THE GIRL HIT THE BOY WHEN HE TOOK THE CANDY
the girl hit the boy when she took the candy

THE GIRL'S FATHER CUT HIMSELF
the girl's father cut herself

THE MAN WHO THE CHILDREN SAW WAS DRAWING A PICTURE
the children saw the man and they were drawing a picture

But score as 5: THE LADY LIKES THE GIRL WHO IS WATCHING THE BABY
the lady likes the girl and she is watching the baby
(Since antecedent of she is ambiguous, give S the benefit of the doubt.)

d. Substitution of words not closely related in meaning or sound, including unrecognizable sounds.

A MAN SAW THE CAR BY THE HOUSE
a man saw the car from the house

e. Omission or addition of possessive pronoun when possession is not obvious.

THE GIRL FOUND HER TOYS
the girl found the toys

THE GIRL HIT THE BOY WHEN HE TOOK THE CANDY
the girl hit the boy when he took his candy

But score as 3: the girl hit the boy when he took her candy

f. Addition of elements which change meaning of sentence.

THE BOY IS THE GIRL'S BROTHER
the little boy is the girl's brother

THE CHILDREN PLAYED A GAME
the children played with a game

Appendix G

Scoring System for the Picture Interview

(With IBM Card Format)

I. System for Coding Verb-Complement Units

Column

1-3 Verb-Complement Unit Number in Transcript

4-6 Sentence Unit Number in Transcript

8-9 Verb-Complement Pattern

1 = S Vi, e.g., He goes (walk, swim, run, breathe, sleep, get in, he got way up there)

2 = S Vl SComp adj, e.g., He looked fierce (turns sour, goes white, looks like [illegible]andy; He feels like swimming)

3 = S Vl SComp noun, e.g., They became men (remain, seem, be, look like--when means seem)

4 = S Vt DO, e.g., He hit her

5 = S Vt IO DO, e.g., He gave her food (Also, He gave food to her)

6 = S Vt DO OComp, e.g., They called him Paul/foolish (obj. comp. with noun or adj.)

7 = S Vt DO OComp, e.g., They elected him president (obj. comp. with noun only)

8 = S Vc SComp adv, e.g., He was outside (Also, loc. prep phrase)

9 = S Vc SComp adj, e.g., It is green (Also, cardinal numbers, superlatives)

10 = S Vc SComp noun, e.g., They are animals (Also, poss. nouns, e.g., mine)

11 = S Vt DO, e.g., He had a dog (No passive transf. possible with lack, befall, have, resemble, etc.)

12 = S Vt DO, e.g., He walks miles (DO of measure)

13 = S Vt (IO) Comp, e.g., It cost me plenty, It lasts me years, It takes me an hour (No passive possible)

14 = S Vt DO DO, e.g., He taught them that (2 DO's possible with ask, tell, teach, strike, and either may appear alone--code only when both present)

15 = S Vt DO, e.g., He looked at the house (inseparable verb-preposition combination) Test by seeing if prep. retained in passive)

16 = S Vt DO, e.g., He looked him up (verb-particle comb.) Particle may appear on either side of object

17 = S Vput DO loc, e.g., He put the ball on the table

Column

18 = S Vcatenative (IO) Comp, e.g., Mary told John to go to school, He tried to get him to say what he was going to do

19 = Cannot determine pattern

11 Expandability

1 = Minimum form, e.g., John hit Mary, He swimming (okay without aux.) Includes predeterminer

2 = Minimum form plus other elements, e.g., John hit Mary on the head (includes any noun modifier except predeterminer)

3 = Deletions in pattern as described—no additions, e.g., Hit Mary (includes both standard and nonstandard deletions such as subject omission in 2nd verb-complement unit when compound verb, copula deletion, etc.)

4 = Deletions plus expansions, e.g., Hit Mary on the head

12 Type of Deletion (if deletion occurs)

1 = Obligatory Deletion, e.g., catenative verbs which do not allow nominal element, subjects of participial and some infinitive phrases, verb deletion in tag questions

2 = Optional Deletion of subject or part of predicate in compounds (Also, deletion of repeated element in correlative clauses and some adverbial clauses)

3 = Optional Deletion of understood words in answering wh- questions (or repeating an answer of this type)

4 = Optional Deletion of Subject with imperative or elliptical questions (See that little thing?)

5 = Copula Deletions

6 = Optional Deletion in question asking clarification of question or imperative

7 = Other Optional Deletions, e.g., nominal element as subject of complement following certain catenative verbs, deletion of DO in relative clauses—The girl the boy saw. . .

8 = Agent deletion in Truncated passive

9 = Other nonstandard deletions

13 Function of Verb Being Coded

1 = Verbal element in dependent adverbial clause

2 = Verbal element in relative clause

3 = Parenthetical clause main verb, e.g., you know, I guess, see

4 = Main verb of sentence

5 = Verbal element in subject constituent

6 = Verbal element in complement constituent

7 = Verbal element in comparative or correlative clauses or other clauses modifying adjectives [he's bigger than I am, He's as tall as a building (is tall), He's afraid that I will come]

Column

8 = Verbal element in phrase acting as object of a preposition (gerundive or infinitive, e.g., waited for the helicopter to take us, waited for him to move down)

9 = Verbal element in adjective phrase (restrictive, participial, or infinitive)

0 = Verbal element in adverbial phrase (nonrestrictive participial or infinitive)

15-16 Connective (code when applicable)

In adverbial clauses: 01 = because; 02 = for; 03 = if; 04 = while; 05 = when; 06 = before; 07 = after; 08 = like; 09 = so, so that; 10 = to, in order to; 11 = where; 12 = until; 13 = everytime

In relative clauses: 01 = who; 02 = that; 03 = which; 04 = what; 06 = like; 07 = where; 95 = optional deletion; 99 = nonstandard deletion

In comparative or correlative or other adjective-modifying clauses: 01 = than; 02 = that; 03 = as

In subject or object constituent clauses and all verbal phrases: 01 = wh- words; 02 = that; 03 = to; 04 = ing; 05 = "quotes"; 06 = -en; 07 = poss.; 08 = how; 10 = for; 11 = if; 13 = like; 14 = because; 34 = (to) + (ing) (get the soldiers to going)--treat other multiple introducers in the same manner as 34; 91 = no introducer-obligatory with some verbs (will you let me ___ smoke?); 92 = no introducer because of compounding of dependent structures; 93 = no introducer because answer to question; 95 = no introducer-optional deletion with verbs such as know (I know ___ he will come); 96 = no introducer because of question asking clarification; 97 = no introducer--optional deletion such as, Go (to) get us; 99 = nonstandard introducer deletion

17 Coordination: Scored in some cases even when connective and, but, etc. are not present, especially when there is subject deletion with the second verb-complement unit. 0 = not coord.; 1 = 1st of 2 or more; 2 = 2nd of 2 or more; 3 = 3rd of 3 or more, etc.

19 Specific Thing Modified

0 = does not modify--is a constituent of a unit; 1 = adverb; 2 = noun whose function cannot be determined; 4 = verb; 5 = subject; 6 = noun complement, DO, or adjective complement; 7 = whole sentence; 8 = adjective, (but not if functions as complement);9 = object of a preposition

20 Kind of Larger Unit Modified:

0 = no modification--is main verb unit; 1 = modifies simple constituent or simple modifier of main verb V-C unit; 2 = modifies element in adverbial clause; 3 = modifies element in relative clause; 4 = modifies element in parenthetical clause; 5 = modifies element in V-C unit acting as subject; 6 = modifies element in V-C unit acting as complement; 7 = modifies element in adjective clause or adjectival V-C phrase; 8 = modifies element in V-C phrase acting as object of a preposition; 9 = modifies element in V-C adverbial phrase

Column

21 Environment of Verb-Complement Unit

1 = Exact or nearly exact repetition of earlier comment (also repetition of specifically requested part of prior utterance) in response to interruption or request to repeat

2 = 1st clause in response to question, request, command from E

3 = 1st clause in response to content statement to E

4 = 1st clause in response to non-directive prompt (including repetition of child's statment)

5 = Not initial after and, so

6 = Not initial after but

7 = Not initial after dependent clause or phrases within the same sentence

8 = Not initial

9 = Initial--Initiated by child

0 = Dependent clause, phrase, or constit. following main verb unit (4)

22 Further Structural Description

1 = Imperative; 2 = Tag question--code on main clause (generalized huh, regular, or nonstandard); 3 = Interrogative yes/no; 4 = Interrogative wh- type; 5 = Unusual Passive (cannot appear in active); 6 = Black English Presentative, e.g., Here go a guitar; 7 = Passive; 8 = Inversion with expletive (only if expletive dropped with re-inversion); 9 = Parenthetical independent clause

23 Negation (code on main clause V--C unit)
1 = Negative form is used; 2 = multiple negation

24 Attachment to Thing Modified
If function of construction is not 4, is the construction attached to thing it modifies or, if a constituent, to the rest of the sentence?
1 = yes; 2 = no; 3 = no, because E requested specific part to be repeated

26 Part of Interview
0 = warm-up; 1 = 1st picture; 2 = 2nd picture; 3 = 3rd picture; 4 = 4th picture; 5 = 5th picture; 6 = 6th picture

28 Is pattern stereotyped one (I don't know, That's all, nothing)
1 = yes; 0 = no

II. System for Coding Noun Phrases

Column

1-3 Verb-complement Unit Number

4-6 Sentence Unit Number

7-9 Noun Phrase Unit Number

11 Noun Phrase Part Number (in case of coordination) (1 if 1 part)

13-17 Type of Noun Phrase (Place 1 in column for each applicable description)

- Col. 9 Simple Non-Prepositional
- 10 Simple Prepositional
- 11 Appositional
- 12 Coordinate
- 13 Relative (phrase or head of clause)

19-27 Pre and Post Head Noun Slots Filled (Record number of words filling slot. Watch for phrases in one slot)

- Col. 15 Predeterminer (all, both, half, only, just, even)
- 16 Determiner (a, the), Possessive nouns
- 17 Numerals (three)
- 18 Specifier/adj., ordinal, superlatives (same, different, other, certain, next, last)
- 19 Adjective
- 20 Noun Modifier of Head Noun
- 21 Head Noun
- 22 Locative Adverbial Post-noun modifier
- 23 Pronominal Appositive
- 24 Post-noun "all," "both"

29 Type of Head Noun

1 = common count

2 = common mass

3 = proper

4 = personal pronoun

5 = possessive pronoun (mine, yours, etc.)

6 = reflexive

7 = relative pronoun

8 = demonstrative

Column

31 Function of Phrase in Sentence (if simple non-prepositional)

1 = subject
2 = direct object
3 = indirect object
4 = verb complement

33 Function of Prepositional Phrase in Sentence (what modify)

1 = subject
2 = direct object
3 = indirect object
4 = verb complement
5 = object of preposition

35 Function of Constituent that appositional phrase stands in apposition to (Same codes as in column 29)

37-41 Characterization of Relative Phrases and Clauses

Col. 33 Head Noun Function (Same codes as column 29)
34 Shared Noun Function (Same codes as column 29)
35 Person or Non-person
36 Type of Relator (1 = who; 2 = that; 3 = which; 4 = what; 5 = with; 6 = that)
37 Restrictive (1) or Nonrestrictive (2)